PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C12Q 1/00, A61K 45/00, 38/17, C07K 14/705, C12N 15/70, 1/21 // (C12N 1/21, C12R 1:19)

A2

(11) International Publication Number: **WO 97/07209**

(43) International Publication Date: 27 February 1997 (27.02.97)

(21) International Application Number: PCT/GB96/01980

(22) International Filing Date: 14 August 1996 (14.08.96)

(30) Priority Data:
9517098.1 21 August 1995 (21.08.95) GB

(71) Applicant *(for all designated States except US)*: IMPERIAL COLLEGE OF SCIENCE TECHNOLOGY AND MEDICINE [GB/GB]; Sherfield Building, Exhibition Road, London SW7 2AZ (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LITTLE, Peter [GB/GB]; Imperial College of Science Technology and Medicine, Biochemistry Dept., Prince Consort Road, London SW7 2BY (GB). HADJANTONAKIS, Anna-Katerina [GB/GB]; Imperial College of Science Technology and Medicine, Biochemistry Dept., Prince Consort Road, London SW7 2BY (GB).

(74) Agents: HARDING, Charles, Thomas et al.; D. Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.

Date of receipt by the International Bureau:
9 September 1996 (09.09.96)

(54) Title: RECEPTOR

(57) Abstract

A method of screening a chemical for subsequent use as a pharmaceutical agent. The method comprises contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats and/or cadherin-like repeats.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

RECEPTOR

The present invention relates to a receptor. In particular the present invention relates to the use of a receptor to screen agents to assess their suitability for subsequent use as pharmaceutical agents, such as therapeutic agents and diagnostic agents.

Receptors are structures that bind chemical stimuli specifically and directly or indirectly transduce a message into the intracellular environment (Watson *et al* 1992 Recombinant DNA Second Edition, Chapter 17, published by Scientific American Books). Some receptors, otherwise known as G-protein coupled receptors (GCRs), are coupled to second messenger systems *via* GTP-binding proteins, otherwise known as G-proteins. The G-proteins connect hormone receptors to adenylate cyclase or other signalling enzymes.

In more detail, the GCRs represent the largest receptor protein class in vertebrates. Typically they are seven-pass transmembrane receptors (see Figure 1). In particular, the GCRs have been shown to be involved in the regulation of a large variety of physiological processes, with many of the genes encoding them mutated in genetic disorders and mutants.

GCRs, can be divided into six families on the basis of their amino acid sequence homologies (similarities) which span across the transmembrane containing region. In this regard, the B family, which is the second largest family, contains the receptors for pituitary adenylate cyclase activating polypeptide, vasoactive intestinal polypeptide (VIP), secretin, growth hormone releasing hormone, diuretic hormone, glucagon, glucagon-like peptide, calcitonin and gastric inhibitory polypeptide.

GCRs can even be placed into functional categories. In this regard, several different types of G-protein coupling have been identified.

For example, the GCR will either interact with an ion channel which is itself a seven-pass transmembrane protein or it can interact with adenylate cyclase, phospholipase C or phospholipase A2, all of which signal to secondary messengers.

The G-protein can either be stimulatory (Gs) or inhibitory (Gi) and can therefore stimulate or inhibit the action of the ion channel or second messenger pathway they are effecting.

All the B family GCRs have been shown to couple to adenylate cyclase *via* a stimulatory G-protein.

In this regard, the present invention provides a new receptor obtainable from animals. The present invention also provides a new use of that receptor.

Thus, according to a first aspect of the present invention there is provided a method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats.

According to a second aspect of the present invention there is provided a method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises cadherin-like repeats.

According to a third aspect of the present invention there is provided a method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats and cadherin-like repeats.

Preferably, the method includes contacting the chemical-receptor complex with a G-protein and determining whether the chemical-receptor complex stimulates the G-protein.

Preferably the receptor resembles or is a GCR.

In the following commentary, the term "receptor according to the present invention" includes the receptor as defined in the above-mentioned aspects of the present invention.

According to a fourth aspect of the present invention there is provided a method of affecting neural development comprising stimulating a receptor with a chemical; wherein the receptor is the receptor according to the present invention. This method can be an *in vitro* or an *in vivo* method.

According to a fifth aspect of the present invention there is provided the use of the receptor according to the present invention to screen chemicals for subsequent use as a pharmaceutical.

According to a sixth aspect of the present invention there is provided a chemical that has been screened by the method of the present invention.

According to a seventh aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats.

According to a eighth aspect of the present invention there is provided a protein comprising the amino acid sequence represented as SEQ. I.D. No. 1, or a fragment, homologue or variant thereof.

According to a ninth aspect of the present invention there is provided a protein comprising the amino acid sequence represented as SEQ. I.D. No. 3, or a fragment, homologue or variant thereof.

According to a tenth aspect of the present invention there is provided a nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 2, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

According to an eleventh aspect of the present invention there is provided a nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 4, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

According to an twelfth aspect of the present invention there is provided a vector capable of expressing the receptor according to the present invention or the protein according to the present invention, or comprising the nucleotide sequence according to the present invention.

According to a thirteenth aspect of the present invention there is provided a construct comprising or capable of expressing any one of the vector according to the present invention, the receptor according to the present invention, the protein according to the present invention, or the nucleotide sequence according to the present invention.

According to a fourteenth aspect of the present invention there is provided a cell, tissue or organ comprising or capable of expressing any one of the construct according to the present invention, the vector according to the present invention, the receptor according to the present invention, the protein according to the present invention, or the nucleotide sequence according to the present invention.

According to a fifteenth aspect of the present invention there is provided an organism comprising or capable of expressing any one of the cell, tissue or organ according to the present invention, the construct according to the present invention, the vector according to the present invention, the receptor according to the present invention, the protein according to the present invention, or the nucleotide sequence according to the present invention.

According to a sixteenth aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

According to a seventeenth aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

According to an eighteenth aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats and cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

According to a nineteenth aspect of the present invention there is provided an assay kit comprising a surface having attached thereto or contained within or on any one of the organism according to the present invention, the cell, tissue or organ according to the present invention, the construct according to the present invention, the vector according to the present invention, the receptor according to the present invention, the protein according to the present invention, or the nucleotide sequence according to the present invention.

Typically the assay kit will comprise a series of titre wells capable of holding a suitable sample of the present invention (e.g. the receptor or the gene coding for same

in cells or in a cell free environment). Preferably, the assay kit comprises a series of titre wells, wherein at least one of which well holds a suitable sample of the present invention (e.g. the receptor or the gene coding for same in cells or in a cell free environment). Optionally, the assay kit may comprise one or more G-proteins.

As mentioned above, if the assay kit of the present invention comprises the receptor of the present invention then that assay kit would be useful for screening chemicals that are capable of interacting with the receptor to form a chemical-receptor complex. With that assay kit, the interaction of the chemical-receptor complex with the G-protein can be observed either directly or indirectly. An example of indirect observation is observing a change (e.g. an increase) in cAMP levels.

Alternatively, if the assay kit of the present invention comprises the nucleotide sequence of the present invention then that assay kit would be useful for screening chemicals for affecting expression of that sequence.

Other aspects of the present invention include the use of the receptor of the present invention to screen for agents that are capable of any one or more of

stimulating the receptor to cause neural cells to divide;

stimulating the receptor to cause neural cells to differentiate;

stimulating the receptor to affect cellular physiology; and

stimulating the receptor for use in the repair of trauma and/or in the treatment of neurodegenerative diseases.

Other aspects of the present invention include the use of the receptor of the present invention for one or more of:

stimulating adenylate cyclase;
increasing cAMP levels; and
promoting neural growth.

These uses can be *in vitro* or *in vivo* uses.

Other aspects of the present invention include NCIMB No. 40766, NCIMB No. 40767, and NCIMB No. 40768.

Further aspects of the present invention include:

a pharmaceutical preparation consisting of or comprising the receptor of the present invention, optionally admixed with or contained within a suitable carrier, diluent or excipient.

a pharmaceutical preparation consisting of or comprising the nucleotide sequence of the present invention, optionally admixed with or contained within a suitable carrier, diluent or excipient.

a pharmaceutical preparation consisting of or comprising a chemical when screened by the method of the present invention, optionally admixed with or contained within a suitable carrier, diluent or excipient.

the use of the receptor according to present invention in the manufacture of a medicament to treat neural disorder.

the use of the nucleotide sequence according to present invention in the manufacture of a medicament to treat neural disorder.

the use of a chemical when screened by the method of the present invention in the manufacture of a medicament to treat neural disorder.

A further aspect of the present invention includes a method of treating a subject in need of, or likely to be in need of, treatment for neural disorder wherein the method comprises administering to the subject a receptor according to the present invention, or a protein expressed by the nucleotide sequence according to the present invention, or a chemical when screened by the method of the present invention.

An additional aspect of the present invention includes a hybrid receptor, and genes coding for the same and vectors etc. comprising same, wherein the hybrid receptor comprises at least a part of the receptor of the present invention and at least a part of another receptor, such as another receptor or even part or all of a G-protein. The hybrid receptor is advantageous as it allows one to affect and/or to tailor the stimulation of the receptor to one or more stimuli.

Preferably the receptor is obtainable from neural tissue.

Preferably the receptor is obtainable from ectodermal cells.

Preferably the receptor is prepared by use of recombinant DNA techniques.

Preferably the receptor is obtainable from deposit NCIMB No. 40766 and/or deposit NCIMB No. 40767.

Alternatively, the receptor is obtainable from deposit NCIMB No. 40768.

Preferably the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 1, or is a fragment, homologue or variant thereof.

Preferably the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 3, or is a fragment, homologue or variant thereof.

Preferably the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 2, or is a fragment, homologue or variant thereof.

Preferably the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 4, or is a fragment, homologue or variant thereof.

Preferably the chemical is screened to determine if it is useful for one or more of:

i. causing neural cells to divide;
ii. causing neural cells to differentiate;
iii. affecting cellular physiology;
iv. repairing trauma;
v. treating neurodegenerative diseases;
vi. stimulating adenylate cyclase production;
vii. increasing cAMP levels;
viii. promoting neural growth.

Preferred embodiments of the present invention therefore include:

i. a receptor capable of interacting with a G-protein and comprising EGF-like repeats and/or cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells;

ii. a receptor comprising the sequence represented as SEQ. I.D. No. 1, or a fragment, homologue or variant thereof capable of being stimulated by a chemical;

iii. a receptor comprising the sequence represented as SEQ. I.D. No. 3, or a fragment, homologue or variant thereof capable of being stimulated by a chemical;

iv. a receptor encoded by the nucleotide sequence represented as SEQ. I.D. No. 2, or a fragment, homologue or variant thereof and wherein the expressed protein is capable of being stimulated by a chemical;

v. a receptor encoded by the nucleotide sequence represented as SEQ. I.D. No. 4, or a fragment, homologue or variant thereof and wherein the expressed protein is capable of being stimulated by a chemical;

vi. a receptor obtainable from deposit NCIMB No. 40766;

vii. a receptor obtainable from deposit NCIMB No. 40767; and

viii. a receptor obtainable from deposit NCIMB No. 40768.

With this aspect of the present invention, the receptor may comprise a plurality of any combination of the features i. to viii. as listed above.

In a highly preferred embodiment the receptor of the present invention is not expressed by the natural genomic DNA sequence when in its natural environment. Thus, highly preferred embodiments include the receptor when prepared by use of recombinant DNA techniques, including the expression of cDNA or a synthetic nucleotide sequence.

Preferably the receptor is expressed by a cDNA sequence that is obtainable from any one or more of deposits NCIMB 40766, NCIMB 40767 and NCIMB 40768.

In addition, or alternatively, preferably the receptor is expressed by a cDNA sequence that comprises sequence shown as SEQ. I.D. No. 2, or a sequence that is complementary thereto, or a sequence having substantial homology therewith, or a sequence containing any suitable codon substitutions but wherein the resultant protein is capable of acting as receptor as herein defined.

In addition, or alternatively, more preferably the receptor is expressed by a cDNA sequence that comprises sequence shown as SEQ. I.D. No. 4, or a sequence that is complementary thereto, or a sequence having substantial homology therewith, or a sequence containing any suitable codon substitutions but wherein the resultant protein is capable of acting as receptor as herein defined.

In a highly preferred embodiment the nucleotide sequence coding for the receptor of the present invention is not in its natural environment and under the control of the promoter with which it is naturally associated which is also in its natural environment.

Thus, highly preferred embodiments include the use of recombinant DNA techniques using for example cDNA or a synthetic nucleotide sequence.

In a highly preferred embodiment the nucleotide sequence coding for the receptor of the present invention is not in its natural environment and under the control of the promoter with which it is naturally associated which is also in its natural environment, wherein the receptor comprises the amino acid sequence shown as SEQ.I.D. No. 1, more preferably SEQ.I.D. No. 3, or variant, fragment or homologue thereof, wherein the nucleotide sequence is a cDNA sequence that is obtainable from any one or more of deposits NCIMB 40766, NCIMB 40767 and NCIMB 40768, and wherein the nucleotide sequence comprises the sequence shown as SEQ. I.D. No. 2, more preferably SEQ. I.D. No. 4, or a sequence that is complementary thereto, or a sequence having substantial homology therewith, or a sequence containing any suitable codon substitutions but wherein the nucleotide sequence codes for a protein that is capable of behaving as a receptor as herein defined.

Other embodiments of the present invention include: a transformed host organism (such as a microorganism, such as *E. coli.*) capable of producing the receptor according to the present invention as a consequence of the introduction of a nucleotide sequence as herein described; a method for preparing the receptor according to the present invention comprising expressing a nucleotide sequence according to the present invention contained in the host organism and isolating the expressed receptor; and a vector (such as a transformed *pBLUESCRIPT* plasmid, *pGEX* plasmid or *pCDNA3* plasmid) incorporating the nucleotide sequence according to the present invention. By way of example, the receptor can be expressed in *E. coli*, baculovirus, yeast or mammalian cell expression systems.

All of the above-mentioned aspects of the present invention optionally include the combination of the receptor of the present invention with a G-protein.

The term "chemical" includes any chemical compound, including nucleotide sequences both in sense and antisense orientation, proteins, enzymes etc. The term also includes a ligand, wherein a ligand is a natural substance that naturally binds to the receptor.

The term "pharmaceutical agent" includes chemicals for use as diagnostic and/or therapeutic purposes. The term also includes pharmaceutical agents for human and/or veterinary applications.

The terms "screen" and "screening" include the use of the receptor according to the present invention to screen agents to assess their suitability for subsequent use as pharmaceutical agents, such as therapeutic agents and diagnostic agents.

The term "receptor" is used in its normal sense as typically meaning a protein that spans the membrane of a cell and that can bind, on its extra-cellular side, a chemical (otherwise known as a ligand). Binding of the chemical causes changes to the receptor that result in a chemical (enzymatic) reaction being initiated on the intra-cellular part of the receptor. These changes are the first part of a signalling chain of actions that

result in some change to the cells physiology. In the case of GCRs, binding of the chemical causes the receptor to effect a G-protein with which it is associated on the inner membrane surface. This disturbance results in changes to the enzymatic state of the G-protein. which then interacts with the signalling system.

The term "G-protein" is used in its normal sense as typically meaning a protein which is associated with a receptor and which is capable of being effected by the receptor. Changes in the receptor (binding of chemicals/ligands) effect the enzymatic state of the G-protein and these changes can effect the interactions of the G-protein with other proteins which are components of a cascade of signalling events. The outcome of the signalling is dependent on the nature of the cell containing receptor and G-protein. In a preferred embodiment, the receptor interacts with a G-protein.

The term "EGF-like repeats" is used in its normal sense as typically meaning a protein sequence similar to the following "consensus" sequence.

$CX_{2-6}CX_{4-6}CX_{5-10}CXCX_{8-22}C$

wherein C is cysteine and X is any amino acid.

Preferably, the receptor of the present invention comprises at least one EGF-like repeat and/or at least one cadherin-like repeat. Typically, the receptor of the present invention has between 1 and/or 36 EGF-like repeats and between 1 to 36 cadherin-like repeats. Preferably, the receptor of the present invention has between 1 and 20 EGF-like repeats and/or between 1 and 20 cadherin-like repeats. Preferably, the receptor of the present invention has between 3 and 10 EGF-like repeats and/or between 3 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has between 4 and 7 EGF-like repeats and/or between 6 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has at least about 6 EGF-like repeats and/or at least about 10 cadherin-like repeats.

Preferably, the receptor of the present invention comprises at least one EGF-like repeat and at least one cadherin-like repeat.

Preferably, the receptor of the present invention has between 1 and 36 EGF-like repeats and between 1 to 36 cadherin-like repeats. Preferably, the receptor of the present invention has between 1 and 20 EGF-like repeats and between 1 and 20 cadherin-like repeats. Preferably, the receptor of the present invention has between 3 and 10 EGF-like repeats and between 3 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has between 4 and 7 EGF-like repeats and between 6 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has at least about 6 EGF-like repeats and at least about 10 cadherin-like repeats.

The term "chemical-receptor complex" includes binding of the chemical to the receptor, such as by hydrogen bonding and/or covalent bonding. The chemical-receptor complex may then interact with a G-protein. Determination of the formation of the chemical-receptor complex can be achieved by conventional techniques. However, it is preferred to determine formation of the chemical-receptor complex by observing the effect of the complex on a G-protein, such as by observing an increase in cAMP levels.

The term "obtainable from" includes directly or indirectly obtaining the receptor. Examples of indirectly obtaining the receptor include expressing the receptor cDNA by means of a suitable expression system.

The terms "variant", "homologue" or "fragment" include any substitution of, variation of, modification of, replacement of, deletion of or addition of one or more amino acid(s)/nucleic acid from or to the sequence providing the resultant protein is capable of behaving as a receptor as herein defined.

The expression "substantial homology", which can be otherwise expressed as "substantial similarity", includes homology with respect to structure and/or nucleotide components, providing the resultant protein is a receptor as herein defined.

With respect to sequence homology (i.e. similarity), preferably there is at least 50 % homology, preferably at least 60% homology, more preferably at least 75% homology, more preferably at least 80% homology, more preferably at least 85% homology, more preferably at least 90% homology, such as at least 95% homology.

The term "complementary" means that the present invention also covers recombinant nucleotide sequences that can hybridise to the recombinant nucleotide sequences.

The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes all or part of the nucleotide sequence according to the present invention directly or indirectly attached to another nucleotide sequence, such as a promoter.

The construct may even contain or express a marker which allows for the selection of the genetic construct in the host into which it has been transferred.

The construct of the present invention preferably comprises a promoter.

The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site in the Jacob-Monod theory of gene expression.

The term "vector" includes an expression vector and a transformation vector. The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression. The term "transformation vector" means a construct capable of being transferred from one species to another - such as from an *E.Coli* plasmid to another host.

The terms "cell", "tissue" and "organ" include cell, tissue and organ *per se* and when within an organism.

The term "organism" in relation to the present invention includes any organism that could comprise the recombinant nucleotide sequence coding for the receptor according to the present invention and/or products obtained therefrom, and/or wherein the recombinant nucleotide sequence according to the present invention can be expressed when present in the organism.

Preferably the organism is a transgenic organism. The term "transgenic organism" in relation to the present invention includes any organism that comprises the recombinant nucleotide sequence coding for the receptor according to the present invention and/or products obtained therefrom, and/or wherein the recombinant nucleotide sequence according to the present invention can be expressed within the organism. Preferably the recombinant nucleotide sequence is incorporated in the genome of the organism.

The term "protein" includes un-modified and modified proteins such as post-translationally modified proteins and glycosylated proteins.

The receptor of the present invention is sometimes referred to as the ME2 protein.

The following samples were deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St Machar Drive, Aberdeen, Scotland, AB2 1RY, United Kingdom, on 18 August 1995:

E. coli Xl-1 blue containing mouse cDNA plasmid ME2(22) which was allocated deposit number NCIMB 40766;

E. coli Xl-1 blue containing mouse cDNA plasmid ME2(78) which was allocated deposit number NCIMB 40767;

E. coli DH1 containing human cosmid ME2HC20 which was allocated deposit number NCIMB 40768.

These deposits are discussed later in the Experimental Section (see Deposits).

Thus, highly preferred embodiments of the present invention include any one of the aforementioned aspects of the present invention but wherein the receptor or the nucleotide sequence coding for same is obtainable from any one or more of deposits NCIMB 40766, NCIMB 40767 and NCIMB 40768.

The present invention will now be described only by way of examples, in which reference shall be made to the following Figures, in which:

Figure 1 is a pictorial representation of a typical GCR;

Figure 2 is a DNA map of the receptor of the present invention;

Figure 3 is a pictorial representation of the receptor of the present invention;

Figure 4 is a schematic representation of expression patterns;

Figure 5 is a restriction map;

Figure 6 presents SEQ ID No. 1;

Figure 7 presents SEQ ID No. 2;

Figure 8 presents SEQ ID No. 3;

Figure 9 presents SEQ ID No. 4; and

Figure 10 presents an analysis of an amino acid sequence (sequence range 1 to 2707).

In some of the following commentary, sections (parts) of the gene coding for the receptor of the present invention are sometimes referred to as ME2(x) wherein "x" represents a particular numbered fraction.

ME2 Genetic Mapping

Our initial studies revealed that the receptor of the present invention is coded by a single copy gene. This single copy gene is conserved in organisms as diverged as human, mice and fruit flies. In particular, the single copy gene maps to human chromosome region 22^{qter} and mouse chromosome 15. In both of these genomes the receptor gene is contained in a region associated with gastrulation and neural mutants and disorders.

Expression Behaviour

To determine the *in vivo* expression of the receptor of the present invention, both reverse transcriptase polymerase chain reaction (RT-PCR) and wholemount *in situ* hybridisation were carried out on wild-type mouse embryos.

RT-PCR analysis showed that the receptor of the present invention is first expressed in the early postimplantation embryo between 4 and 6 days post coitum (dpc), then continues until adulthood.

The embryonic expression of the receptor of the present invention precedes the start of gastrulation, the event which results in the generation of the three germ layers of the developing embryo. Prior to this, the embryo does not contain neural tissue. Embryonic expression of the gene coding for the receptor of the present invention correlates with cells of ectodermal origin, which go on to form the nervous system.

For ease of reference, Figure 4 is a schematic representation of the expression patterns of the receptor of the present invention in the developing central nervous system, in particular in the developing spinal cord (Figure 4(A)) and the developing hindbrain (Figure 4(B)). In this regard, interesting features of this dynamic expression pattern include the delineation of segments in the developing hindbrain and neural tube. In the hindbrain novel sub-rhombomeric expression was observed. In the neural tube, initially the transcripts are ubiquitous and then resolve into 5,4 and finally 2 dorsoventrally restricted domains, one in the roof plate and one in the floor plate. Gene expression is highly localised and persists throughout development, with adult transcripts localised to the brain and eye.

Expression Discussion

Observing the pattern of expression of the receptor of the present invention indicates that it may be involved in the control of neural development. In this regard, the receptor is expressed almost exclusively in neural tissue (which is discussed in more detail later). In particular, expression precedes the first obvious neural structures in the developing embryo and the pattern of subsequent expression is complex. In later embryos, expression around the ventricle of the brain is significant since this is believed to be the region that contains the neural stem cells.

Further observations revealed detection of the receptor of the present invention peri-ventricularly in adult brains. This particular pattern of synthesis is therefore under complex spatial and temporal controls and is the period in which the nervous system is proliferating most rapidly.

Hence, the expression evidence strongly suggests that the receptor of the present invention might play an important part in the "control machinery" of neural development (see later discussion).

Isolation Of The Receptor Of The Present Invention

In some of the following commentary, sections (parts) of the gene coding for the receptor of the present invention are sometimes referred to as ME2(x) wherein "x" represents a particular numbered fraction. The isolation of the complete the receptor of the present invention coding sequence is shown in Figure 5.

In more detail, in step 1 (Figure 5) a mouse 8.5 dpc whole embryo cDNA was screened using a human cDNA clone (16FB2). 16FB2 was originally isolated from a human fetal brain cDNA library by hybridisation with human cosmid ZnFP16 as described in Hoovers *et al.*, (Genomics 10 254-263).

ME2(2) was then isolated from the initial screening of the mouse cDNA library (Figure 5). Extensive sequence analysis of both ME2(2) and 16FB2 has shown that they have sequence homology in a G-rich region in the 3' untranslated region.

Furthermore, complete nucleotide sequencing of ME2(2) showed that it had no homology to any sequences in the publicly accessible DNA databases.

ME2(2) was then used immediately in whole mount *in situ* expression analyses and produced the striking expression pattern of the receptor of the present invention. The remainder of the gene sequence coding for the receptor of the present invention was then isolated as follows.

In step 2 (Figure 5) ME2(2) was used to re-screen the mouse 8.5dpc cDNA library leading to the isolation of 6 different clones, the largest of which being ME2(22). Sequence analysis and database searches with the 3.2kb ME2(22) sequence identified a large open reading frame whose predicted amino acid sequence had homology to the family B group of G-protein coupled receptors. In this regard, ME2(22) covers the region from the polyA tail to trans-membrane region IV.

In step 3 (Figure 5) the 5' EcoRI to *Pst*I fragment of ME2(19) was used to rescreen which led to the isolation of 3 further cDNA clones.

In step 4 (Figure 5) a primer PLKH20 corresponding to sequence 3527 to 3541 in the sequence shown as SEQ.I.D. No. 2 and derived from ME2(42) was used to isolate the H1 fragment using the RACE method (Frohman, M.A. (1993) Methods in Enzymol. 218 340-56). The DNA sequence of H1 was used to identify the DNA from 3295 to 3312 (again see SEQ. I.D. No. 2) which was used to make a new primer, PLKH26, for RACE analysis which gave rise R12 in step 5 (Figure 5).

In step 6 (Figure 5), R12 was used to screen the cDNA library which gave rise to, amongst other clones, the cDNA clone ME2(78) which extends almost to the 5'end.

The minimal set of cDNA clones that defines all of the receptor of the present invention is ME2(78) and ME2(22) (see Figure 2 and Figure 5).

DNA Analysis

The DNA sequence of the receptor of the present invention was established entirely using published methods.

In particular, the sequencing methodology used was the Sanger technique (Sanger *et al.*, 1977 Proc. Natl. Acad. Sci. USA. 12 5463-7). The sequencing kits used were supplied by Pharmacia Biotech and the manufacturer's protocols were followed throughout.

DNA was sequenced either by analysis of cloned molecules using sequencing primers specific for vector sequences and sequencing into the cDNA, or by synthesising specific primers, obtained from conventional commercial synthesis companies, and using these to establish DNA sequence directly from internal parts of the cloned cDNA molecule.

The sequence data obtained is shown in the attached Figures as SEQ. I.D. No. 2 (see Figure 7) and SEQ. I.D. No. 4 (see Figure 9). A map of the DNA sequence is represented in Figure 2.

Re-isolation Of The Receptor Of The Present Invention

The receptor of the present invention was re-isolated by using PCR techniques.

Since the gene coding for the receptor of the present invention is over 4.5 kb it is preferable to isolate a cDNA containing the receptor coding regions by a multi-step process, rather than by a one step process using RT-PCR to isolate the whole cDNA. Hence, by using the complete coding sequence for the receptor of the present invention it is possible to isolate a series of cDNA fragments that can then be ligated.

In this regard, primers flanking pairs of unique restriction enzyme sites were used to amplify individual regions and subsequent restriction enzyme digestion and ligation to generate the complete sequence. Details of this approach are given below and in Figure 2.

Fragment 1: Primer pair(3') PLKH31+(5') PLKH52
Amplification products are digested with *BspHI*.

Fragment 2: Primer pair(3') PLKH47+(5')PLKH59c
Amplification products are digested with *BspHI*

Primer	Sequence (5'-3')	Position
PLKH59C	5' CAG CGG GGA CTA CTG CGA GAC TGA AAT	1-27
PLKH47	5' AGC TTG TCG AAG ATG TCA AC	2675-2694
PLKH52	5' ATC TTA CAG CAT GAG AGC CGC C	2414-2435
PLKH31	5' GGT AAT GAC ACA GTC ACT GGC ATG	4856-4879

Fragments 1 and 2 were independently amplified from mouse embryonic or adult brain reverse transcribed cDNA under standard PCR conditions. The amplification products were then subsequently directly restriction endonuclease digested with BspHI to give ragged ends. Fragments 1 and 2 were then ligated to each other and then cloned into a T vector.

Construction Of A Complete cDNA Clone

Construction of a complete cDNA clone for the gene for the receptor of the present invention was as follows.

In particular, construction of a complete cDNA clone for the gene for the receptor of the present invention was achieved using the protocol detailed in the previous section.

In more detail, the method used relied on the minimal set of cDNA clones obtained from the mouse 8.5dpc libraries mentioned above (see also Figure 2 and Figure 5).

The minimal set of clones ME2(22) and ME2(78) have *EcoR*I linkers and are cloned into pBluescript plasmid vector. Since there is no *EcoR*I site in the ME2 transcript, construction of a complete cDNA clone will be done by *EcoR*I + *Avr*II digestion of ME2(22), isolation of the 2.6kb fragment and ligation of this to the 4.3kb *EcoR*I plus *Avr*II fragment from ME2(78): this product is cloned into the appropriate *EcoR*I digested vector.

Deposits

As mentioned above three deposits have been made in accordance with the Budapest Treaty. In this regard, NCIMB 40766 is an *E. coli* XI-1 blue host with a *pBluescript* SK+ vector containing fragment ME2(22) - i.e. nucleotides 3657 to 6794 (see SEQ.I.D.No. 2).

NCIMB 40767 is an *E. coli* Xl-1 blue host with a *pBluescript* SK+ vector containing fragment ME2(78) - i.e. nucleotides 1 to 4813 (see SEQ.I.D. No. 2).

NCIMB 40768 is a recombinant cosmid containing the main part of the human receptor gene according to the present invention. In this regard, the cosmid vector is *pCos*2EMBL and the host cell *E. coli* DH1. The human DNA derives from a region of human chromosome 22^{qter} and contains parts of the human receptor of the present invention gene including parts of the 7TM region but not extending further 5' than this.

In order to prepare a full length cDNA clone from NCIMB 40766 and NCIMB 40767, the appropriate cDNA fractions can be excised by use of suitable restriction enzymes, isolated and then ligated. The full length cDNA can then be inserted into any suitable expression system and subsequently expressed by suitable means. The pBluescript SK+ plasmids can be recovered from bacterial cells grown in L-broth containing 100 (μg/ml Ampicillin using routine methods detailed in Sambrook *et al.*, (1989 Molecular cloning - A laboratory manual. Second edition. Cold Spring harbor laboratory press.). Then the receptor cDNA fragments may be isolated subsequent to their excision with *EcoR*I.

Likewise, the DNA from NCIMB 40768 and/or fragments thereof can be excised by use of suitable restriction enzymes, isolated, inserted into any suitable expression system and subsequently expressed by suitable means. The DNA can be recovered by growing the bacteria in L-broth supplemented with 30 (μg/ml Kanamycin and recovering the DNA according to routine methods detailed in Sambrook *et al.*, (1989 Molecular cloning - A laboratory manual. Second edition. Cold Spring harbor laboratory press). The human DNA fragments can be resolved by cleavage with almost any 6-base recognition enzyme.

Amino acid analysis

The amino acid sequence data are listed in the attached sequence listings as SEQ. I.D. No. 1 (see Figure 6) and SEQ. I.D. No. 3 (see Figure 8).

Amino acid analysis of the receptor of the present invention reveals that it appears to be a large membrane spanning receptor having an unusual structure. This structure is pictorially shown in Figure 3.

In slightly more detail, the C-terminal region appears to have the structure of a 7-pass transmembrane receptor related to the B family of G-protein coupled receptors (GCRs). Thus it is believed that the receptor of the present invention is a new protein.

Further amino acid analysis reveals that the receptor of the present invention contains EGF-like repeats. In this regard, towards the N-terminus (extracellular) of the receptor there are a number of EGF-like repeats (see Figure 3). One of these EGF-like repeats is ~300 amino acids away from the C-terminal sequence. The EGF-like repeats are shown in Figure 10 (marked "EGF 1" etc.). Divergent EGF-like repeats are also marked.

A large number of molecules containing EGF-like repeats have been identified in both vertebrates and invertebrates. These molecules include, for example, blood clotting factors and proteins that are required for correct embryonic development.

Examples of proteins that are required for correct embryonic development, which molecules have primarily been characterised in invertebrates, include fibropellin, a cell coat protein of sea urchins, glp-1 and lag-12 proteins required for inductive interactions in the nematode worm *C. elegans*, and a number of *Drosophila* proteins including Crumbs, which is required for establishing epithelial cell polarity. Additional examples include Notch, Delta and Serrate proteins, which are required for neurogenesis, and Slit, which is a protein involved in axonal pathfinding. Notch and

its ligands Delta and Serrate are involved in cell-cell signalling that determines adjacent cell fates; though this signal is not directly mitogenic.

To date, there have been reports of some isolated vertebrate proteins that have some homology (similarity) to the invertebrate proteins. For example, three Notch genes, a single Delta, Jagged and Delta-like have been identified to date.

Other vertebrate proteins that have been isolated are EMR-1 (Baud et al 1995 Genomics 26 334-344) and CD97 (Genebank accession no. X84700). However, there is no mention of possible utility of such proteins, let alone mention of pattern of expression/synthesis.

Further amino acid analysis reveals that the receptor of the present invention contains cadherin-like repeats. These cadherin-like repeats are shown in Figure 10 (marked as "CD 1" etc.). Cadherin-like repeats have been implicated in protein-protein interactions (Geiger and Ayalon 1992 Ann Rev Cell Biol 8 307-332).

Some of the transmembrane portions of the receptor of the present invention are shown in Figure 10 (marked as "TM 1" etc.).

When the amino acid sequence of the receptor of the present invention is compared with the amino acid sequences of known proteins that are required for correct embryonic development it is observed that there is some sequence homology, though this is less than 80%. More importantly, however, in distinction to the receptor of the present invention those proteins are all single-pass transmembrane proteins with the cluster of EGF-like repeats in their extracellular domains. In contrast, the receptor of the present invention has a seven-pass transmembrane topology, similar to a GCR.

Accordingly, as there have been no reports in the literature for a receptor found in neural tissue that is capable of interacting with a G-protein but, in addition, having EGF-like repeats so the receptor of the present invention is novel.

Functions Of The Receptor Of The Present Invention

The EGF-like repeats and the 7 pass transmembrane (7TM) structures of the receptor of the present invention suggest important functions and utilities for the receptor. In this regard, the EGF-like repeats will bind a ligand which may, as in the case of Notch/Delta, be a protein attached to another cell, or it may be free, as in the case of blood clotting factors. In either case, binding of a ligand will cause the cytoplasmic region to signal to the cellular machinery, *via* a G-protein. Since the 7TM region is most similar to the family B GCRs regions, it is likely that it will signal *via* a stimulatory G-protein which stimulates adenylate cyclase and causes cAMP levels to increase as all family B receptors operate in this fashion.

Likewise, the cadherin-like repeats of the receptor of the present invention suggest important functions and utilities for the receptor.

Without wishing to be bound by theory, it is believed that the outcome of this signalling, based upon known examples of GCRs, could be due to one or more of the following effects:

1. Stimulation of the receptor of the present invention might cause neural cells to divide.

2. Stimulation of the receptor of the present invention might cause neural cells to differentiate.

3. Stimulation of the receptor of the present invention might cause changes to cellular physiology.

Uses

Based on the above-mentioned functions of the receptor of the present invention, it is clear that the receptor can be used in a number of useful utilities. Some of those utilities are now presented.

1. The receptor of the present invention could be a therapeutic target. In this regard, if the ligand or a modified ligand can be defined, then artificial treatment of neural tissue with this ligand could trigger the receptor of the present invention to signal. This signal would then trigger the normal response of the receptor of the present invention, which would be a way of modulating the growth, function or properties of brain cells that express the receptor of the present invention. This would therefore have an application in the repair of trauma and in treatment of neurodegenerative diseases, such as Parkinson's disease and Alzheimer's disease.

2. The receptor of the present invention can be used as a test reagent to identify the ligand. This can be done for example in artificial test systems where the receptor of the present invention is expressed from a recombinant expression vector in cells that otherwise do not express the receptor of the present invention. These cells can then be used to test for binding of the ligand by studying cAMP level changes upon treatment of cells with proteins or other cells.

3. The nucleotide sequence coding for the receptor of the present invention gene could be used to modify the behaviour of cells by transgenesis. Potential areas of application include the modification of cells used for transplantation treatments of degenerative diseases and the modification of whole animals by normal transgenic procedures. Both of these applications would result in cells with modified growth potential.

Screening Protocol

Two screening protocols are now presented.

The first is based on Lutz *et al.* (1993. FEBS Letters 334, 3-8). In outline, a ME2 expression vector (for example pcDNA-1) is constructed and then introduced into Cos-7 cells by transfection. This results in the cells expressing ME2 and, because of the biological properties of G-proteins, a G-protein becomes naturally associated with ME2 on the cell surface. The cells are then treated *in vitro* with proteins, chemicals, other cells (intact or broken up). Then one assays for changes to cAMP levels which are caused by ME2 binding a ligand and signalling via the G-protein to alter cAMP levels. If changes are seen, this implies that the ligand is or is contained in, in the substance that was treated with cells. This is the assay of choice for purification of the ligand.

The second method is based upon Cheng & Flanagan (1994, Cell 79 157-168). In this regard, one synthesises in *E. coli* and isolates the N-terminal fragment of ME2 (N terminus of ME2 to the membrane entry point in fig 3) which has been fused at this point to the enzyme alkaline phosphatase. This hybrid protein binds to its normal ligand. Binding is then detected by looking for the alkaline phosphatase dragged along at its end. This could be used to isolate cDNA clones containing the normal ME2 ligand using exactly the methods detailed in Cheng & Flanagan (1994, Cell 79 157-168).

The present invention therefore relates to a novel receptor and a novel use of that receptor.

The nature of the receptor of the present invention, its spatiotemporally restricted expression, coupled with the evolutionary conservation of the receptor gene suggests that the receptor of the present invention plays a role in the determination of ectodermal cell types within the developing embryo. This has important consequences

as it enables possible pharmaceutical agents to be screened to see if they stimulate the receptor and if so then those agents could be used to promote neural growth. In addition, the receptor can be inserted (such as by way of transplantation or by way of transgenesis of the coding gene) into a subject either in need of treatment or to develop *in vivo* screening techniques.

Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope of the invention.

CLAIMS

1. A method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor: and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats.

2. A method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor: and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises cadherin-like repeats.

3. A method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats and cadherin-like repeats.

4. A method according to any one of claims 1 to 3 wherein the receptor is obtainable from neural tissue.

5. A method according to any one of claims 1 to 4 wherein the receptor is obtainable from ectodermal cells.

6. A method according to any one of the preceding claims wherein the receptor is prepared by use of recombinant DNA techniques.

7. A method according to any one of the preceding claims wherein the receptor is obtainable from deposit NCIMB No. 40766 and/or deposit NCIMB No. 40767.

8. A method according to any one of claims 1 to 6 wherein the receptor is obtainable from deposit NCIMB No. 40768.

9. A method according to any one of the preceding claims wherein the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 1, or is a fragment, homologue or variant thereof.

10. A method according to any one of the preceding claims wherein the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 3, or is a fragment, homologue or variant thereof.

11. A method according to any one of the preceding claims wherein the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 2, or is a fragment, homologue or variant thereof.

12. A method according to any one of the preceding claims wherein the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 4, or is a fragment, homologue or variant thereof.

13. A method according to any one of the preceding claims wherein the method includes contacting the chemical-receptor complex with a G-protein and determining whether the chemical-receptor complex stimulates the G-protein.

14. A method according to any one of claims 1 to 13 wherein the chemical is screened to determine if it is useful for one or more of:

i. causing neural cells to divide;
ii. causing neural cells to differentiate;
iii. affecting cellular physiology;
iv. repairing trauma;
v. treating neurodegenerative diseases;
vi. stimulating adenylate cyclase production;
vii. increasing cAMP levels;
viii. promoting neural growth.

15. A method of affecting neural development comprising stimulating a receptor with a chemical; wherein the receptor is the receptor as defined in any one of claims 1 to 14.

16. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor to cause neural cells to divide.

17. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor to cause neural cells to differentiate.

18. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor to affect cellular physiology.

19. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor for use in the repair of trauma and/or in the treatment of neurodegenerative diseases.

20. Use of a receptor as defined in any one of claims 1 to 14 to stimulate adenylate cyclase.

21. Use of a receptor as defined in any one of claims 1 to 14 to increase cAMP levels.

22. Use of a receptor as defined in any one of claims 1 to 14 to promote neural growth.

23. Use of a receptor as defined in any one of claims 1 to 14 to screen chemicals for subsequent use as a pharmaceutical.

24. A chemical when screened by the method according to any one of claims 1 to 15.

25. A receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

26. A receptor capable of interacting with a G-protein and wherein the receptor comprises cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

27. A receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats and cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

28. A receptor according to any one of claims 25 to 27 wherein the receptor is the receptor as defined in any one of claims 4 to 14.

29. A protein comprising the amino acid sequence represented as SEQ. I.D. No. 1, or a fragment, homologue or variant thereof.

30. A protein comprising the amino acid sequence represented as SEQ. I.D. No. 3, or a fragment, homologue or variant thereof.

31. A nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 2, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

32. A nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 4, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

33. A vector capable of expressing the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

34. A construct comprising or capable of expressing any one of the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

35. A cell, tissue or organ comprising or capable of expressing any one of the construct of claim 34, the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

36. An organism comprising or capable of expressing any one of the cell, tissue or organ of claim 35, the construct of claim 34, the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

37. An assay kit comprising a surface having attached thereto or contained within or on any one of an organism according to claim 36, the cell, tissue or organ of claim 35, the construct of claim 34, the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

38. NCIMB No. 40766.

39. NCIMB No. 40767.

40. NCIMB No. 40768.

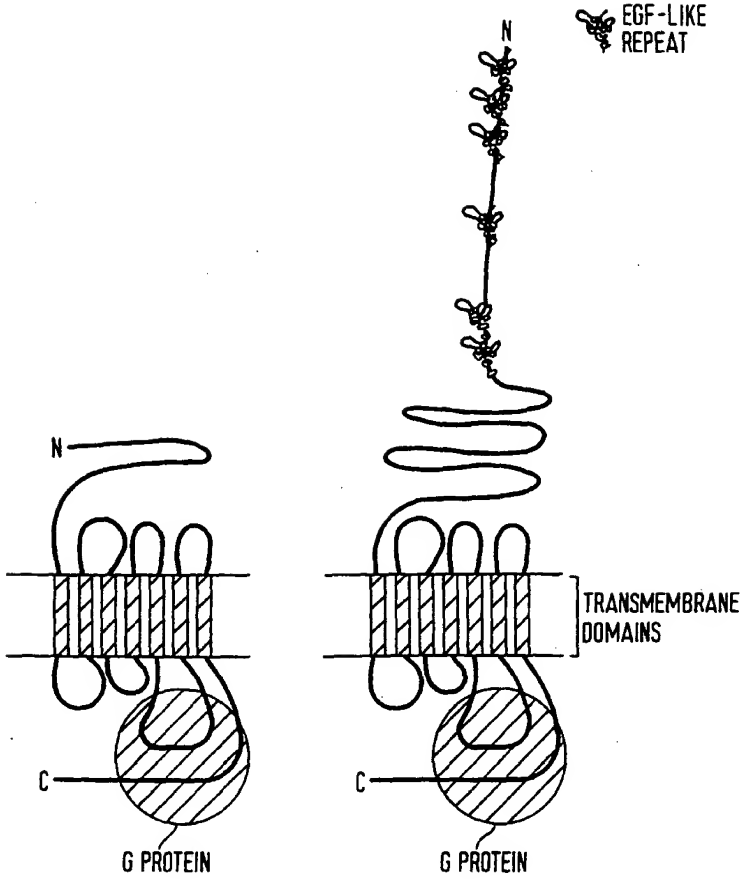
41. A method substantially as described herein.

42. A use substantially as described herein.

43. A receptor substantially as described herein.

44. An amino acid sequence substantially as described herein.

45. A nucleotide sequence substantially as described herein.



TYPICAL FAMILY B
RECEPTOR STRUCTURE

FIG. 1

ME2
RECEPTOR STRUCTURE

FIG. 3

FIG. 2

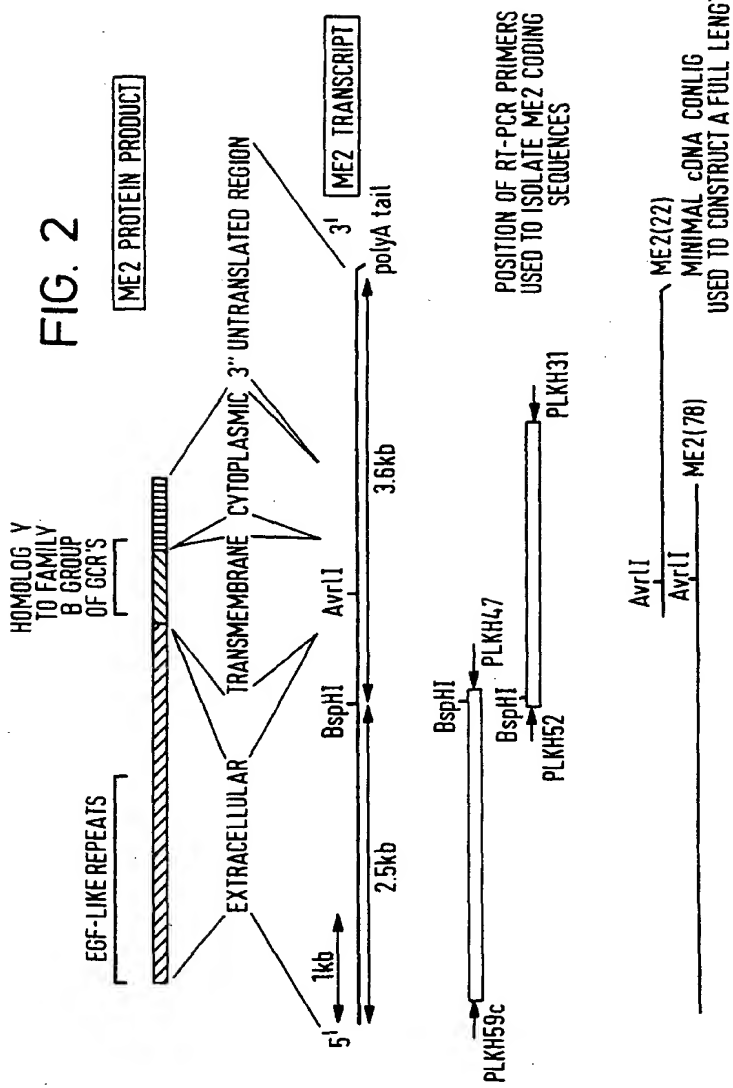
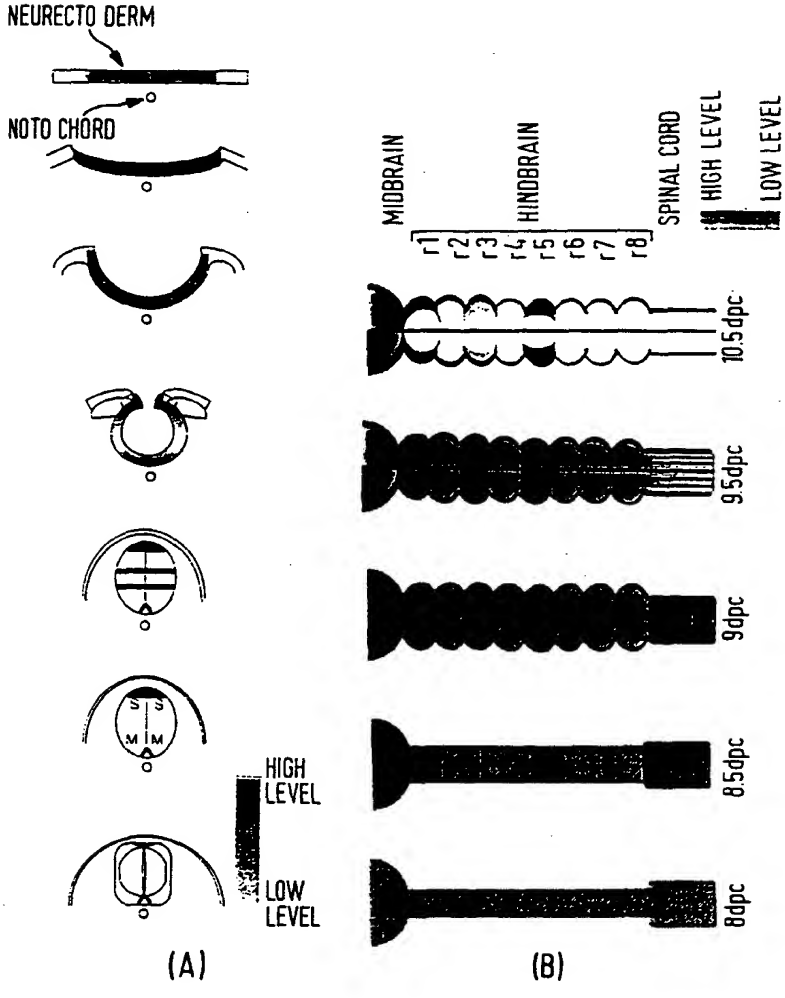

ME2 PROTEIN PRODUCT
ME2 TRANSCRIPT
HOMOLOG Y TO FAMILY B GROUP OF GCR'S
EGF-LIKE REPEATS
EXTRACELLULAR
TRANSMEMBRANE
CYTOPLASMIC
3" UNTRANSLATED REGION
5'
3'
1kb
2.5kb
3.6kb
BspHI
AvrII
polyA tail
PLKH59c
PLKH47
PLKH52
PLKH31
POSITION OF RT-PCR PRIMERS USED TO ISOLATE ME2 CODING SEQUENCES
ME2(22)
ME2(78)
MINIMAL cDNA CONLIG USED TO CONSTRUCT A FULL LENGTH ME2 CLONE

FIG. 4

EXPRESSION OF ME2 IN THE DEVELOPING MOUSE SPINAL CORD (A) AND HINDBRAIN (B)



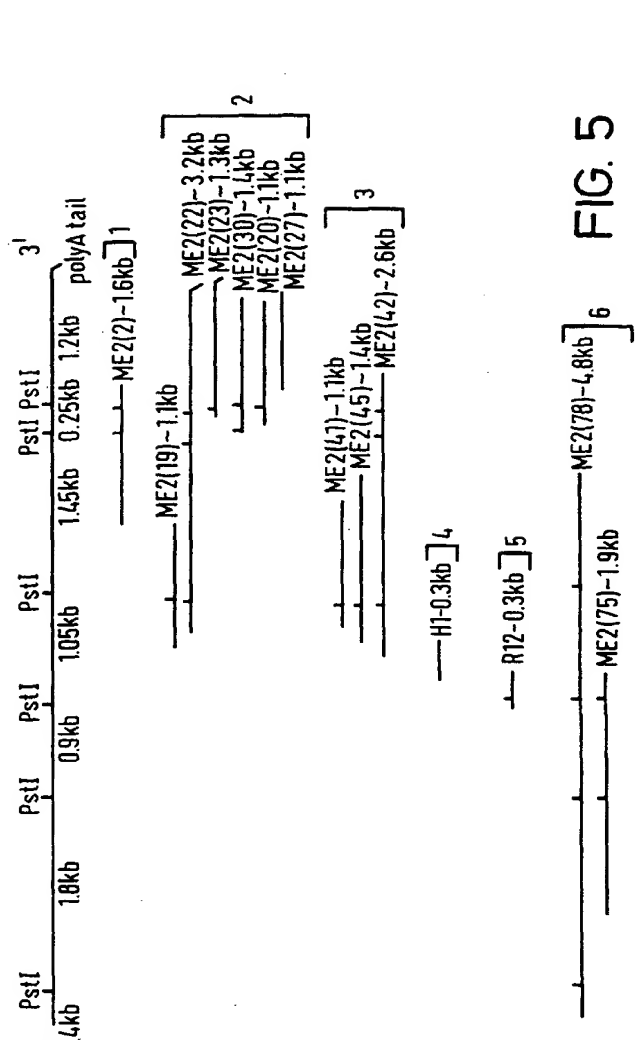

ME2 TRANSCRIPT WITH POSITIONS OF ALL cDNA CLONES
5'
PstI
PstI
PstI
PstI
PstI
3'
0.4kb
1.8kb
0.9kb
1.05kb
1.45kb
0.25kb
1.2kb
polyA tail
ME2(2)-1.6kb
1
ME2(19)-1.1kb
ME2(22)-3.2kb
ME2(23)-1.3kb
ME2(30)-1.4kb
ME2(20)-1.1kb
ME2(27)-1.1kb
2
ME2(41)-1.1kb
ME2(45)-1.4kb
ME2(42)-2.6kb
3
H1-0.3kb
4
R12-0.3kb
5
ME2(78)-4.8kb
ME2(75)-1.9kb
6

FIG. 5

AMINO ACID SEQUENCE OF ME2 REGION COVERED BY ME2(78) AND ME2(22)

1573 RESIDUES

GDYCETEIDLCYSNPCGANGGCRSREGGYTCECFEDFTGEHCQVNVRSGRCASGVCK
NGGTCVNLLIGGFHCVCPPGEYEHPYCEVSTRSFPPQSFVTFRGLRQRFHFTVSLAF
ATQDRNALLLYNGRFNEKHDFIALEIVEEQLQLTFSAGETTTTVTPQVPGGVSDGRW
HSVLVQYYNKPNIGHLGLPHGPSGEKVAVVTVDDCDAAVAVHFGSYVGNYSCAAQGT
QSGSKKSLDLTGPLLLGGVPNLPEDFPVHSRQFVGCMRNLSIDGRIVDMAAFIANNG
TRAGCASQRNFCDGTSCQNGGTCVNRWNTYLCECPLRFGGKNCEQAMPHPQRFTGES
VVLWSDLDITISVPWYLGLMFRTRKEDGVLMEATAGTSSRLHLQILNSYIRFEVSYG
PSDVASMQLSKSRITDGGWHHLLIELRSAKEGKDIKYLAVMTLDYGMDQSTVQIGNQ
LPGLKMRTIVIGGVTEDKVSVRHGFRGCMQGVRMGESSTNIATLNMNDALKVRVKDG
CDVEDPCASSPCPPHRPCRDTWDSYSCICDRGYLEKKCVDACLLNPCKHVGSLCALP
NTPRGYSCECGPGHYGQYCENKVDLPCPKGWWGNRCVAPVTVLSAKALIPTATRPMA
SARRITTSPQPRIVAFPVTVSPRSHSRACDMDTGQCACKPGVIGRQCNRCDNPFAEV
TSLGCEVIYNGCPRAFEAGIWWPQMKFGQPAAVLCPKGSVGNAVRHCSGEKGWLPPE
LFNCTSGSFVDLKALNEKLNRNETRMDGNRSLRLAKALRNATQGNSTLFGNDVRTAY
QLLARILQHESRQQGFDLAATREANFHEDVVHTGSALLAPATEASWEQIQRSEAGAA
QLLRHFEAYFSNVARNVKRTYLRPFVIVTANMILAVDIFDKLNFTGAQVPRFEDIQE
ELPRELESSVSFPADTFKPPEKKEGPVVRLTNRRTTPLTAQPEPRAERETSSSRRRR
HPDEPGQFAVALVVIYRTLGQLLPEHYDPDHRSLRLPNRPVINTPVVSAMVYSEGTP
LPSSLQRPILVEFSLLETEERSKPVCVFWNHSLDTGGTGGWSAKGCELLSRNRTHVT
CQCSHSASCAVLMDISRREHGEVLPLKIITYAALSLSLVALLVAFVLLSLVRTLRSN
LHSIPQEPIHALFFSQLIFMVGINQTENPFLCTVVAILLHYVSMGTFAWTLVENLHV
YRMLTEVRNIDTGPMAFYHVVGWGIPAIVTGLAVGLDPQGYGNPDFCWLSLQDTLIW
SFAGPVGTVIIINTVIFVLSAKVSCQRKHHYYERKGVVSMLRTAFLLLLLVTATWLL
GLLAVNSDTLSFHYLFAAFSCLQGIFVLLFYCVANREVRKHLRAVLAGKKLQLDDSA
TTRATLLTRSLNCNNTYSEGSRHAPHRPGQSTASLDSTTRDEGVQKLSVSSGPARGN
HGEPDASFIPRNSKKAHGPDSDSDSELSLDEHSSSYASSHTSDSEDDGGEAEDKWNP
AGGPAHSTPKADALANHVPAGWPDESLAGSDSEELDTEPHLKVRPRSAWSYTGRRRA
ITVATGPLTRKVGSWPSQWPCLAASPRSSGKAS*

SEQ ID NO. 1

FIGURE 6

COMPLETE NUCLEIC ACID SEQUENCE OF ME2 REGION COVERED BY ME2(78) AND ME2(22) cDNA CLONES

NUCLEOTIDES 1 - 6791

CCGGGGACTACTGCGAGACTGAAATTGATCTTTGCTACTCCAATCCGTGCGGGGCCA
ATGGCGGCTGCCGGAGCCGTGAGGGTGGCTACACTTGTGAGTGCTTCGAGGACTTCA
CTGGGGAGCATTGCCAGGTGAACGTTCGCTCAGGCCGCTGTGCCAGCGGAGTATGCA
AAAACGGGGGCACCTGCGTGAACCTGCTCATTGGAGGCTTCCACTGTGTGTGCCCGC
CCGGCGAGTATGAGCATCCCTACTGTGAAGTGAGCACCAGGAGCTTCCCACCCCAGT
CCTTCGTTACCTTCCGAGGCCTGCGGCAACGCTTCCACTTCACCGTCTCCCTGGCGT
TTGCCACCCAGGACAGGAATGCCCTGCTGCTCTACAATGGCCGCTTCAATGAGAAGC
ACGACTTCATCGCCCTGGAGATTGTGGAGGAGCAGCTGCAGCTCACGTTCTCGGCAG
GTGAGACCACAACCACGGTGACACCGCAGGTTCCTGGAGGTGTGAGCGATGGGCGGT
GGCATTCGGTGCTGGTGCAGTACTACAACAAGCCCAACATTGGCCACCTGGGCCTGC
CCCACGGGCCGTCTGGAGAGAAGGTGGCTGTGGTGACTGTGGATGACTGTGACGCAG
CGGTGGCCGTGCACTTTGGAAGTTACGTGGGGAACTACAGCTGCGCTGCCCAGGGCA
CTCAGAGCGGCTCCAAGAAGTCACTGGATCTGACTGGTCCTCTGCTTCTGGGTGGTG
TCCCCAACCTGCCAGAAGACTTCCCCGTGCACAGCCGTCAGTTTGTGGGATGCATGC
GAAACCTGTCCATCGATGGCCGGATTGTGGACATGGCTGCGTTTATTGCCAACAATG
GTACCAGGGCAGGCTGTGCTTCTCAGAGGAACTTCTGCGATGGGACCTCATGCCAGA
ACGGGGGCACCTGTGTGAACAGGTGGAACACGTACTTATGTGAGTGCCCGCTCCGCT
TTGGTGGAAAGAACTGTGAACAAGCTATGCCACACCCTCAGCGCTTCACTGGTGAGA
GCGTCGTGTTGTGGAGTGACCTTGACATCACCATCTCTGTGCCTTGGTACCTGGGGC
TCATGTTCCGGACCCGGAAGGAGGATGGTGTGCTGATGGAAGCCACAGCTGGCACGT
CTTCCAGGCTCCATCTCCAGATTCTCAACAGCTACATCCGCTTTGAGGTCTCCTACG
GCCCCTCTGACGTGGCATCCATGCAGCTGTCCAAGTCCCGGATAACTGACGGGGGGT
GGCATCACCTGCTCATAGAACTGAGGAGTGCCAAGGAGGGCAAGGACATCAAATACC
TGGCAGTCATGACCTTGGACTATGGGATGGACCAGAGCACAGTGCAGATTGGGAATC
AGCTTCCTGGGTTGAAGATGCGGACTATTGTCATCGGAGGTGTGACCGAGGACAAGG
TCTCTGTCCGCCATGGTTTCCGAGGCTGTATGCAGGGAGTGAGGATGGGAGAGAGCT
CCACCAACATTGCCACCCTGAACATGAATGACGCCCTCAAGGTCAGGGTGAAGGACG
GCTGTGATGTGGAGGACCCATGTGCCTCAAGCCCCTGCCCTCCCCATAGACCCTGCC
GTGACACATGGGACAGCTACTCCTGCATCTGTGACAGAGGGTACTTGGAAAAAAAGT
GTGTGGATGCGTGTCTCCTGAACCCCTGCAAGCACGTTGGCAGCCTGTGTGCGCTCC
CCAACACTCCTCGAGGCTACTCCTGCGAGTGCGGACCCGGCCACTATGGGCAGTACT
GTGAGAACAAAGTCGACCTTCCGTGCCCCAAAGGCTGGTGGGGGAACCGGTGTGTGG
CCCCTGTCACTGTGCTGTCAGCCAAGGCTTTGATCCCGACTGCAACAAGACCAATGG
CCAGTGCAAGGAGAATTACTACAAGCCCCCAGCCCAGGATCGTTGCCTTCCCTGTGA
CTGTTTCCCCCCGCTCCCACAGCCGTGCCTGCGACATGGACACTGGGCAGTGTGCCT
GCAAGCCTGGTGTCATCGGCCGTCAGTGCAACCGCTGTGATAATCCTTTCGCGGAGG
TCACCTCGCTCGGCTGTGAAGTGATCTACAATGGGTGTCCCAGAGCATTTGAGGCTG
GCATCTGGTGGCCACAGATGAAATTTGGGCAGCCAGCAGCGGTGCTATGCCCCAAAG
GATCCGTGGGTAACGCAGTCCGGCACTGCAGTGGGGAGAAGGGCTGGCTTCCCCCAG
AGCTCTTCAACTGCACCTCTGGCTCCTTTGTGGACCTCAAGGCCTTGAACGAGAAAC
TGAACCGCAACGAGACAAGAATGGACGGGAACCGGTCCCTGCGGCTGGCAAAGGCTC
TGAGGAACGCCACGCAGGGGAACAGCACCCTCTTTGGCAATGATGTGCGCACGGCCT
ACCAGCTTCTGGCCCGCATCTTACAGCATGAGAGCCGCCAGCAGGGCTTTGACCTGG
CAGCCACCCGAGAGGCTAATTTTCATGAGGATGTCGTCCATACAGGCAGCGCCCTCC
TGGCCCCAGCTACAGAGGCATCGTGGGAACAGATCCAGCGAAGCGAGGCTGGTGCAG
CGCAGCTACTGAGGCATTTCGAGGCATACTTCAGCAACGTGGCACGAAATGTGAAGA
GGACCTATCTGAGGCCCTTCGTCATCGTCACCGCCAACATGATTCTTGCAGTTGACA
TCTTCGACAAGCTGAACTTCACGGGTGCCCAGGTGCCAAGGTTTGAGGACATTCAGG
AAGAGCTCCCAAGGGAGCTGGAGTCCTCCGTGTCCTTCCCAGCTGACACCTTCAAGC

SEQ ID NO 2

FIGURE 7

CACCAGAGAAAAAAGAAGGCCCCGTGGTGAGGCTGACCAACCGGAGGACTACCCCAC
TCACGGCACAACCAGAGCCCAGGGCTGAGAGGGAAACCTCATCCAGCAGACGGAGGA
GACACCCCGATGAGCCTGGACAGTTTGCTGTTGCCCTGGTTGTCATTTACCGGACCC
TGGGTCAGCTGCTGCCTGAACACTATGACCCCGACCATCGCAGCCTCCGACTGCCTA
ACCGGCCTGTCATCAACACCCCCGTGGTGAGTGCTATGGTGTACAGTGAGGGAACCC
CACTCCCCAGCTCTCTGCAGAGGCCTATCCTGGTGGAGTTCTCCCTGTTGGAGACGG
AGGAACGAAGCAAACCTGTCTGTGTATTCTGGAACCACTCCCTCGACACTGGTGGGA
CTGGAGGGTGGTCAGCCAAGGGCTGTGAACTTCTGTCGAGGAACCGGACCCACGTCA
CTTGCCAGTGCAGCCATTCGGCCAGCTGCGCGGTGCTCATGGACATTTCCAGACGTG
AGCACGGGGAGGTTCTGCCCCTGAAGATCATCACCTATGCCGCCCTGTCCTTGTCTT
TGGTGGCCCTCCTGGTGGCCTTTGTCCTTCTCTCGCTCGTTCGGACACTGCGCTCCA
ACCTGCACAGCATCCCCACAAGAACCTATCCACGCTCTGTTCTTCTCCCAGCTCATCT
TCATGGTCGGCATCAACCAGACTGAGAACCCGTTTCTCTGCACAGTGGTCGCCATCC
TCCTGCACTACGTCTCCATGGGCACCTTCGCCTGGACCCTTGTGGAGAACTTGCATG
TCTACCGCATGCTGACAGAAGTGCGCAACATCGACACTGGGCCCATGGCGTTCTACC
ACGTGGTGGGCTGGGGCATCCCTGCCATTGTCACAGGACTGGCTGTTGGCTTGGACC
CTCAGGGCTATGGAAACCCTGACTTCTGCTGGCTGTCCCTTCAGGATACCCTGATTT
GGAGCTTTGCTGGGCCTGTCGGAACGGTTATAATCATCAACACAGTCATCTTTGTCC
TGTCTGCAAAGGTTTCCTGCCAAAGAAAGCACCATTATTATGAAAGAAAGGGGGTTG
TCTCCATGCTGAGGACGGCCTTCCTCCTGCTGCTGCTCGTCACTGCCACCTGGCTGC
TGGGACTGCTGGCGGTCAACAGTGACACTCTTAGCTTTCACTACCTCTTTGCTGCCT
TCAGCTGCTTGCAGGGCATCTTTGTCCTCCTGTTCTACTGCGTGGCCAACAGGGAGG
TGCGGAAGCACCTGAGGGCGGTGCTGGCAGGGAAGAAGCTGCAGCTGGATGACTCGG
CCACCACTCGGGCCACTCTGCTAACGCGCTCCCTCAACTGCAACAACACCTACAGCG
AAGGGTCCAGACATGCTCCGCACCGCCCTGGGCAGTCCACAGCCTCTCTGGACAGTA
CCACCAGGGATGAAGGGTCCAGAAACTCAGTGTGTCCTCTGGCCCAGCCCGTGGTA
ACCATGGAGAACCAGATGCATCCTTCATCCCTAGGAACTCCAAAAAAGCTCACGGCC
CTGACTCTGACTCTGACAGTGAGCTGTCCCTGGACGAGCACAGTAGTTCCTACGCCT
CTTCACACACATCGGACAGCGAGGATGATGGCGGAGAGGCTGAAGACAAATGGAATC
CGGCTGGGGGCCCCGCCCATAGCACCCCAAAAGCAGATGCTCTGGCCAACCACGTCC
CAGCTGGCTGGCCCGACGAGAGCCTGGCTGGGAGTGACAGTGAGGAGTTGGACACTG
AGCCCCACCTGAAGGTGAGACCAAGGTCAGCGTGGAGTTACACCGGCAGGCGCAGGG
CAATCACTGTGGCGACCGGCCCTCTGACCCGGAAAGTGGGGTCCTGGCCAAGCCAGT
GGCCGTGCTTAGCAGCCAGCCCCAGGAGCAGCGGAAAGGCATCCTGAAAAACAAAGT
CACCTACCCGCCGGCCATTGCCAGAGCAGCCACTGAAGTCCCGGCTGCGAGAGAAGC
TGGCTGATTGTGAGCAGAGCCCCACATCCTCCCGCACATCCTCCCTTGGCTCTGGCG
ATGGTGTCCATGCCACTGACTGTGTCATTACCATCAAGACTCCGAGGAGGGAGCCAG
GCCGTGAGCATCTCAATGGGGTGGCCATGAATGTACGCACAGGGAGTGCCAGGCCAA
CGGTTCTGACTCAGAGAAACCATGAGGCACAGTCAACCCCACAGACTGCCGGTCAAG
CCCTCAGACCTTGAAGCCTGCCTGGGACTGCTGCCTATGGGACAAGCAGGCACCTTG
TGTGAGGGTCCCTGCCATAGCAGCTGGCTCTACGCAGACCGTCCAGACGGGAAGCCC
TTGACCCTCATAGGAGCTCAGGGCCCAGACTCTGACAAAGTGCCAAAGCCACAGATG
TCTCCAGAGGGAGACGTGGACTTCATTAAGGCTGGCATAGCTCCGTCCTTAGACTGA
AGACAGAATCCAAACCATGCTGTACAAGAGGCCATTGAGCCAGAGCTGGACTTGGTG
AATCATTGTACCGGGCCCTTCAACTGTCCCGCAGGCCTCTCCTTGTGGTACAAGCCC
ATCACCACCAGCCTAGCGGTGCCTCTGCAACGGCAACCCTGGGTTTTAAATTGCTGT
CTAAAAATGTAAATAGATATAAATCTCTCCCTGGACTTGGGAGAAGATGGGAGCTGT
GTATGCTTTACACTGCTTTGACTCTGCAGCCACTGGAGAGCCATGAAATGGCATCTC
ACTCTATTGCCCAAGGAAGCTTGCACAGTTGTACTTGAATCTGGAATGAGTCAACTC
AGCTGGTCCCAGTGCCCAGGTAGGGAGCATATGGGCTGTGAAGTTGACACAGCTTCT
GCAGCTGCTTCCGGCCAGGGAGTGTGGGTGTTCCGCTGGCCTTTGGAGCACACGGCA
CATGTGCTGCGTATGTGTGCCTCAGCTCCATGGATCCAGGTCAGGGCTCACCTGTGA
GGAGTGGGCGCCATAGCTATGATATGAAACTCTGACCACCCTGCCACCCCCACGCCC
CCCCCCCCCCCCCGCTCCTGATGCAGTGAGGCGACTCTGGAGCCTTTCCCAGTCAGC

FIGURE 7 CONTINUED

CCAGTGCGCAGCGGGAGGACTTCTGCTCCTGCTCCCATAAGCTCTCAGAGTGCTCAT
GTTCCTGATTCTGAGGGAGCCCAGAGGTCTTCCAACGCTGTCACGGTCCTTCCTGCC
AATGTTGTACTTTTTCAGACTCCAGGGCTCTCTAGGCAGGCAGGTGCTCCACCCCTG
AAATCTGACCAATGACATCACTTTGCTTCAAATGACCAATTGTGCAAAGAAACAAAG
CCAGAGTGCCCGTTTTCAATGGTTACCACATTCTTTTTGGAATCTCACACCAAGGAC
CTGTGACCAGCCACTGAGAGCCACGGTGCAGCCAAGGCCAAGGGATGGGAGCCTGGA
GTTACTCAGACACGTTTACTTCAGCATGGACTGTTGTCTGAATCAGGTCCCCAAAGT
ACATGGGGTGCACAGTCGCTCGGAATGGAGAACCTCAGGCGAGGCGGTCAAAGGCCA
GGACTTCTCACCAAGCTGTGCTCTCAGAAGTGCACAGGACTGCTAGTGACTGACTGG
TGAGATGAAAGCCTACAAGACATGGCCTGGGGTCACGCACTGCCAAGAGGCTGACGG
GAGGCCCAGGTAGCCCAAGGATGGCAAGGATACAGAGTGACCTAGCACAGGGGAGCT
TCAGTCCCAGGTGGTACAGCACCGTGACAACTCCGCAACCCACGCCACCTCAGAAGG
TGAAGTTTTTGATTCGATCACAACTATTAGCAAAACAAACCCTGTCAGTTTTAAACT
GTTTTTCTGACCTAAGACTTTCCTTGACCGTAAGACATGGAGATTTTAACAGGTGTA
TTTATTACTGTTCAGCACTGGATGGCAACACAGGTGAGATGATGCATCTATAATAAA
TTAAGATTTTTGGATTTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAA

FIGURE 7 CONTINUED

ME2 PROTEIN 1-1798 AMINO ACIDS

NSGLMEKLKQIEEQTKKAQQELEEQTRRALELEQERKRAGTAVIELRAHDPDEGDAGRLSYQMEALFDER
SNGYFLIDAATGAVTTARSLDRETKDTHVLKVSAVDHGSPRRSAATYLTVTVSDTNDHSPVFEQSEYRER
IRENLEVGYEVLTIRATDGDAPSNANMRYRLLEGAGGVFEIDARSGVVRTRAVVDREEAAEYQLLVEAND
QGRNPGPLSASATVHIVVEDENDNYPQFSEKRYVVQVPEDVAVNTAVLRVQATDRDQGQNAAIHYSIVSG
NLKGQFYLHSLSGSLDVINPLDFEAIREYTLRIKAQDGGRPPLINSSGLVSVQVLDVNDNAPIFVSSPFQ
AAVLENVPLGHSVLHIQAVDADAGENARLQYRLVDTASTIVGGSSVDSENPASAPDFPFQIHNSSGWITV
CAELDREEVEHYSFGVEAVDHGSPAMSSSASVSITVLDVNDNDPMFTQPVYELRLNEDAAVGSSVLTLRA
RDRDANSVITYQLTGGNTRNRFALSSQSGGGLITLALPLDYKQERQYVLAVTASDGTRSHTAQVFINVTD
ANTHRPVFQSSHYTVSVSEDRPVGTSIATISATDEDTGENARITYVLEDPVPQFRIDPDTGTIYTMTELD
YEDQAAYTLAITAQDNGIPQKSDTTSLEILILDANDNAPRFLRDFYQGSVFEDAPPSTSVLQVSATDRDS
GPNGRLLYTFQGGDDGDGDFYIEPTSGVIRTQRRLDRENVAVYNLWALAVDRGSPNPLSASVGIQVSVLD
INDNPPVFEKDELELFVEENSPVGSVVARIRANDPDEGPNAQIIYQIVEGNVPEVFQLDLLSGDLRALVE
LDFEVRRDYMLVVQATSAPLVSRATVHIRLLDQNDNPPELPDFQILFNNYVTNKSNSFPSGVIGRIPAHD
PDLSDSLNYTFLQGNELSLLLLDPATGELQLSRDLDNNRPLEALMEVSVSDGIHSVTALCTLRVTIITDD
MLTNSITVRLENMSQEKFLSPLLSLFVEGVATVLSTTKDDIFVFNIQNDTDVSSNILNVTFSALLPGGTR
GRFFPSEDLQEQIYLNRTLLTTISAQRVLPFDDNICLREPCENYMKCVSVLRFDSSAPFISSTTVLFRPI
HPITGLRCRCPPGFTGDYCETEIDLCYSNPCGANGGCRSREGGYTCECFEDFTGEHCQVNVRSGRCASGV
CKNGGTCVNLLIGGFHCVCPPGEYEHPYCEVSTRSFPPQSFVTFRGLRQRFHFTVSLAFATQDRNALLLY
NGRFNEKHDFIALEIVEEQLQLTFSAGETTTTVTPQVPGGVSDGRWHSVLVQYYNKPNIGHLGLPHGPSG
EKVAVVTVDDCDAAVAVHFGSYVGNYSCAAQGTQSGSKKSLDLTGPLLLGGVPNLPEDFPVHSRQFVGCM
RNLSIDGRIVDMAAFIANNGTRAGCASQRNFCDGTSCQNGGTCVNRWNTYLCECPLRFGGKNCEQAMPHP
QRFTGESVVLWSDLDITISVPWYLGLMFRTRKEDGVLMEATAGTSSRLHLQILNSYIRFEVSYGPSDVAS
MQLSKSRITDGGWHHLLIELRSAKEGKDIKYLAVMTLDYGMDQSTVQIGNQLPGLKMRTIVIGGVTEDKV
SVRHGFRGCMQGVRMGESSTNIATLNMNDALKVRVKDGCDVEDPCASSPCPPHRPCRDTWDSYSCICDRG
YLEKKCVDACLLNPCKHVGSLCALPNTPRGYSCECGPGHYGQYCENKVDLPCPKGWWGNRCVAPVTVLSA
KALIPTATRPMASARRITTSPQPRIVAFPVTVSPRSHSRACDMDTGQCACKPGVIGRQCNRCDNPFAEVT
SLGCEVIYNGCPRAFEAGIWWPQMKFGQPAAVLCPKGSVGNAVRHCSGEKGWLPPELFNCTSGSFVDLKA
LNEKLNRNETRMDGNRSLRLAKALRNATQGNSTLFGNDVRTAYQLLARILQHESRQQGFDLAATREANFH
EDVVHTGSALLAPATEASWEQIQRSEAGAAQLLRHFEAYFSNVARNVKRTYLRPFVIVTANMILAVDIFD
KLNFTGAQVPRFEDIQEELPRELESSVSFPADTFKPPEKKEGPVVRLTNRRTTPLTAQPEPRAERETSSS
RRRRHPDEPGQFAVALVVIYRTLGQLLPEHYDPDHRSLRLPNRPVINTPVVSAMVYSEGTPLPSSLQRPI
LVEFSLLETEERSKPVCVFWNHSLDTGGTGGWSAKGCELLSRNRTHVTCQCSHSASCAVLMDISRREHGE
VLPLKIITYAALSLSLVALLVAFVLLSLVRTLRSNLHSIPQEPIHALFFSQLIFMVGINQTENPFLCTVV
AILLHYVSMGTFAWTLVENLHVYRMLTEVRNIDTGPMAFYHVVGWGIPAIVTGLAVGLDPQGYGNPDFCW
LSLQDTLIWSFAGPVGTVIIINTVIFVLSAKVSCQRKHHYYERKGVVSMLRTAFLLLLLVTATWLLGLLA
VNSDTLSFHYLFAAFSCLQGIFVLLFYCVANREVRKHLRAVLAGKKLQLDDSATTRATLLTRSLNCNNTY
SEGSRHAPHRPGQSTASLDSTTRDEGVQKLSVSSGPARGNHGEPDASFIPRNSKKAHGPDSDSDSELSLD
EHSSSYASSHTSDSEDDGGEAEDKWNPAGGPAHSTPKADALANHVPAGWPDESLAGSDSEELDTEPHLKV
RPRSAWSYTGRRRAITVATGPLTRKVGSWPSQWPCLAASPRSSGKAS*

SEQ ID NO. 3

FIGURE 8

ME2 DNA SEQUENCE 1 - 8210

GAATTCCGGGCTGATGGAGAAGCTGAAGCAGATTGAGGAGCAGACTAAGAAGGCTCAGC
AAGAGCTGGAAGAGCAGACCCGCAGGGCCCTAGAACTTGAGCAGGAACGGAAGCGTGCG
GGCACTGCGGTCATCGAACTGCGCGCGCACGACCCAGACGAAGGCGATGCAGGACGCCT
TAGCTACCAGATGGAGGCGCTGTTCGATGAGCGCTCTAATGGCTACTTCCTCATCGATG
CCGCCACGGGTGCAGTGACCACAGCCCGCTCCCTGGACCGGGAGACCAAGGACACTCAT
GTACTCAAAGTTAGTGCTGTGGACCACGGCTCCCCGAGGCGCTCAGCTGCCACCTACCT
CACCGTAACTGTCAGTGACACTAACGACCACAGCCCAGTCTTTGAGCAGTCTGAGTATC
GAGAGCGAATTCGAGAAAACCTGGAGGTGGGCTATGAGGTTCTGACCATCCGTGCCACC
GACGGGGATGCCCCTTCCAACGCAAACATGCGCTATCGTCTGCTGGAGGGCGCAGGTGG
TGTCTTTGAGATAGACGCACGATCAGGTGTCGTGCGCACACGGGCTGTGGTGGACCGTG
AGGAGGCGGCTGAGTACCAGCTGCTGGTGGAGGCCAATGACCAGGGTCGCAATCCAGGC
CCACTCAGTGCCTCAGCCACCGTCCACATAGTGGTAGAAGACGAGAATGACAACTACCC
CCAGTTCAGTGAGAAGCGCTATGTGGTTCAAGTCCCAGAAGACGTAGCCGTCAACACGG
CTGTGCTTCGAGTCCAGGCCACTGACCGGGACCAGGGGCAGAATGCAGCCATACACTAC
AGCATCGTTAGTGGCAACCTGAAGGGTCAGTTCTACCTGCATTCGCTTAGTGGGAGCCT
GGATGTTATCAACCCGCTGGACTTCGAAGCCATCCGGGAATACACCCTGCGCATCAAAG
CCCAAGATGGGGGCCGGCCTCCTCTCATTAATTCCTCAGGACTGGTCTCGGTGCAGGTG
TTAGATGTGAACGACAATGCGCCCATCTTTGTGAGCAGCCCCTTTCAGGCTGCCGTGCT
AGAGAATGTGCCCCTCGGCCACTCAGTCCTGCACATCCAAGCGGTGGACGCAGATGCAG
GGGAGAACGCCAGGCTGCAGTACCGTCTAGTGGACACAGCCTCCACTATCGTGGGGGGC
AGCAGTGTCGACTCTGAGAACCCTGCCTCTGCCCCAGACTTCCCCTTCCAAATCCACAA
CAGCTCCGGTTGGATTACTGTGTGCGCGGAGCTGGACCGTGAGGAGGTGGAACACTATA
GCTTTGGAGTAGAAGCAGTGGACCATGGCTCACCAGCCATGAGCTCCTCTGCCAGCGTG
TCCATCACAGTGCTGGATGTAAATGATAACGACCCCATGTTCACGCAGCCTGTGTATGA
GCTGCGTCTGAATGAGGATGCGGCTGTCGGGAGCAGCGTGCTGACCCTCAGGGCCCGAG
ACCGTGATGCCAATAGTGTGATCACCTACCAGCTGACGGGTGGGAACACCCGCAACCGC
TTCGCACTCAGCAGCCAGAGCGGCGGTGGCCTTATCACCTTGGCACTGCCCCTGGACTA
CAAGCAGGAACGGCAGTATGTGCTGGCTGTGACCGCGTCCGATGGCACGCGTTCACACA
CCGCGCAGGTCTTTATCAACGTTACAGATGCCAACACCCACAGGCCGGTTTTCCAGAGT
TCCCACTACACGGTCAGTGTGAGTGAAGACCGGCCCGTGGGCACCTCCATCGCTACCAT
CAGTGCCACGGATGAGGATACGGGTGAGAACGCCCGCATCACCTATGTGCTAGAGGATC
CCGTACCACAGTTCCGCATTGACCCGGACACTGGCACCATTTATACCATGACGGAACTG
GACTATGAGGACCAGGCTGCCTACACGCTGGCCATCACGGCTCAGGACAATGGCATTCC
TCAGAAGTCAGACACTACCTCTCTGGAGATCCTTATCCTCGACGCCAATGACAACGCGC
CCAGGTTCCTGCGAGATTTCTACCAGGGTTCTGTTTTCGAGGATGCCCCCCCATCTACC
AGTGTCCTCCAGGTCTCTGCTACAGACCGTGACTCAGGCCCTAATGGCCGCCTCCTGTA
CACTTTCCAGGGTGGGGATGATGGAGATGGAGATTTCTACATTGAGCCCACGTCTGGTG
TGATCCGTACCCAGCGCCGGCTGGACAGAGAGAATGTGGCCGTGTACAACCTTTGGGCT
CTCGCTGTGGATCGGGGCAGCCCGAATCCCCTCAGTGCGTCAGTGGGAATTCAGGTGAG
TGTGTTGGACATTAACGACAACCCCCCAGTGTTTGAGAAGGACGAGCTGGAGCTGTTTG
TGGAAGAGAACAGCCCTGTGGGTTCAGTGGTAGCAAGAATAAGGGCCAACGACCCGGAC
GAAGGTCCGAATGCTCAGATCATTTATCAGATCGTGGAGGGCAATGTGCCCGAGGTCTT
CCAGCTGGACCTACTGAGTGGTGACCTCCGTGCCCTGGTCGAGTTGGATTTTGAGGTCC
GGAGGGACTATATGTTGGTGGTGCAGGCCACGTCTGCTCCTCTGGTAAGCCGGGCCACC
GTGCACATCCGTCTCCTGGACCAGAATGACAACCCACCGGAGTTGCCTGACTTCCAGAT
CCTTTTCAACAACTATGTCACCAATAAATCCAACAGCTTCCCCAGTGGTGTGATCGGCC
GCATCCCAGCCCACGACCCTGACCTATCTGACAGCCTCAATTACACCTTTCTGCAAGGC
AACGAGCTGAGCCTGCTGCTGCTGGATCCCGCCACAGGAGAGTTGCAGCTCAGCCGGGA
TCTGGACAACAACCGGCCACTGGAGGCGCTCATGGAGGTGTCTGTGTCAGATGGTATCC
ACAGCGTCACCGCTCTCTGCACTCTGCGCGTGACCATCATTACAGATGACATGCTGACC
AACAGCATCACTGTCCGCCTGGAGAACATGTCGCAGGAGAAGTTCCTGTCCCCGCTGCT
GTCCCTCTTTGTAGAAGGGGTGGCCACAGTACTGTCCACCACCAAGGATGACATCTTCG
TCTTCAACATCCAGAACGACACGGACGTCAGCTCCAACATCCTGAACGTGACTTTCTCG
GCACTGCTCCCCGGTGGCACCCGTGGCCGGTTCTTCCCGTCAGAGGACCTGCAGGAGCA
GATCTACCTGAACCGGACACTGCTCACCACCATCTCCGCCCAGCGTGTGCTGCCCTTTG
ATGACAACATCTGCCTGAGGGAGCCCTGCGAGAACTACATGAAGTGCGTGTCCGTGCTT

SEQ IN NO.4

FIG 9

AGGTTTGACAGTTCGGCACCCTTCATTAGTTCCACCACGGTGCTCTTCCGGCCTATCCA
TCCCATCACGGGCCTGCGCTGCCGCTGCCCGCCGGGTTTCACCGGGGACTACTGCGAGA
CTGAAATTGATCTTTGCTACTCCAATCCGTGCGGGGCCAATGGCGGCTGCCGGAGCCGT
GAGGGTGGCTACACTTGTGAGTGCTTCGAGGACTTCACTGGGGAGCATTGCCAGGTGAA
CGTTCGCTCAGGCCGCTGTGCCAGCGGAGTATGCAAAAACGGGGGCACCTGCGTGAACC
TGCTCATTGGAGGCTTCCACTGTGTGTGCCCGCCCGGCGAGTATGAGCATCCCTACTGT
GAAGTGAGCACCAGGAGCTTCCCACCCCAGTCCTTCGTTACCTTCCGAGGCCTGCGGCA
ACGCTTCCACTTCACCGTCTCCCTGGCGTTTGCCACCCAGGACAGGAATGCCCTGCTGC
TCTACAATGGCCGCTTCAATGAGAAGCACGACTTCATCGCCCTGGAGATTGTGGAGGAG
CAGCTGCAGCTCACGTTCTCGGCAGGTGAGACCACAACCACGGTGACACCGCAGGTTCC
TGGAGGTGTGAGCGATGGGCGGTGGCATTCGGTGCTGGTGCAGTACTACAACAAGCCCA
ACATTGGCCACCTGGGCCTGCCCCACGGGCCGTCTGGAGAGAAGGTGGCTGTGGTGACT
GTGGATGACTGTGACGCAGCGGTGGCCGTGCACTTTGGAAGTTACGTGGGGAACTACAG
CTGCGCTGCCCAGGGCACTCAGAGCGGCTCCAAGAAGTCACTGGATCTGACTGGTCCTC
TGCTTCTGGGTGGTGTCCCCAACCTGCCAGAAGACTTCCCCGTGCACAGCCGTCAGTTT
GTGGGATGCATGCGAAACCTGTCCATCGATGGCCGGATTGTGGACATGGCTGCGTTTAT
TGCCAACAATGGTACCAGGGCAGGCTGTGCTTCTCAGAGGAACTTCTGCGATGGGACCT
CATGCCAGAACGGGGGCACCTGTGTGAACAGGTGGAACACGTACTTATGTGAGTGCCCG
CTCCGCTTTGGTGGAAAGAACTGTGAACAAGCTATGCCACACCCTCAGCGCTTCACTGG
TGAGAGCGTCGTGTTGTGGAGTGACCTTGACATCACCATCTCTGTGCCTTGGTACCTGG
GGCTCATGTTCCGGACCCGGAAGGAGGATGGTGTGCTGATGGAAGCCACAGCTGGCACG
TCTTCCAGGCTCCATCTCCAGATTCTCAACAGCTACATCCGCTTTGAGGTCTCCTACGG
CCCCTCTGACGTGGCATCCATGCAGCTGTCCAAGTCCCGGATAACTGACGGGGGGTGGC
ATCACCTGCTCATAGAACTGAGGAGTGCCAAGGAGGGCAAGGACATCAAATACCTGGCA
GTCATGACCTTGGACTATGGGATGGACCAGAGCACAGTGCAGATTGGGAATCAGCTTCC
TGGGTTGAAGATGCGGACTATTGTCATCGGAGGTGTGACCGAGGACAAGGTCTCTGTCC
GCCATGGTTTCCGAGGCTGTATGCAGGGAGTGAGGATGGGAGAGAGCTCCACCAACATT
GCCACCCTGAACATGAATGACGCCCTCAAGGTCAGGGTGAAGGACGGCTGTGATGTGGA
GGACCCATGTGCCTCAAGCCCCTGCCCTCCCCATAGACCCTGCCGTGACACATGGGACA
GCTACTCCTGCATCTGTGACAGAGGGTACTTGGAAAAAAAGTGTGTGGATGCGTGTCTC
CTGAACCCCTGCAAGCACGTTGGCAGCCTGTGTGCGCTCCCCAACACTCCTCGAGGCTA
CTCCTGCGAGTGCGGACCCGGCCACTATGGGCAGTACTGTGAGAACAAAGTCGACCTTC
CGTGCCCCAAAGGCTGGTGGGGGAACCGGTGTGTGGCCCCTGTCACTGTGCTGTCAGCC
AAGGCTTTGATCCCGACTGCAACAAGACCAATGGCCAGTGCAAGGAGAATTACTACAAG
CCCCCAGCCCAGGATCGTTGCCTTCCCTGTGACTGTTTCCCCCCGCTCCCACAGCCGTG
CCTGCGACATGGACACTGGGCAGTGTGCCTGCAAGCCTGGTGTCATCGGCCGTCAGTGC
AACCGCTGTGATAATCCTTTCGCGGAGGTCACCTCGCTCGGCTGTGAAGTGATCTACAA
TGGGTGTCCCAGAGCATTTGAGGCTGGCATCTGGTGGCCACAGATGAAATTTGGGCAGC
CAGCAGCGGTGCTATGCCCCAAAGGATCCGTGGGTAACGCAGTCCGGCACTGCAGTGGG
GAGAAGGGCTGGCTTCCCCCAGAGCTCTTCAACTGCACCTCTGGCTCCTTTGTGGACCT
CAAGGCCTTGAACGAGAAACTGAACCGCAACGAGACAAGAATGGACGGGAACCGGTCCC
TGCGGCTGGCAAAGGCTCTGAGGAACGCCACGCAGGGGAACAGCACCCTCTTTGGCAAT
GATGTGCGCACGGCCTACCAGCTTCTGGCCCGCATCTTACAGCATGAGAGCCGCCAGCA
GGGCTTTGACCTGGCAGCCACCCGAGAGGCTAATTTTCATGAGGATGTCGTCCATACAG
GCAGCGCCCTCCTGGCCCCAGCTACAGAGGCATCGTGGGAACAGATCCAGCGAAGCGAG
GCTGGTGCAGCGCAGCTACTGAGGCATTTCGAGGCATACTTCAGCAACGTGGCACGAAA
TGTGAAGAGGACCTATCTGAGGCCCTTCGTCATCGTCACCGCCAACATGATTCTTGCAG
TTGACATCTTCGACAAGCTGAACTTCACGGGTGCCCAGGTGCCAAGGTTTGAGGACATT
CAGGAAGAGCTCCCAAGGGAGCTGGAGTCCTCCGTGTCCTTCCCAGCTGACACCTTCAA
GCCACCAGAGAAAAAAGAAGGCCCCGTGGTGAGGCTGACCAACCGGAGGACTACCCCAC
TCACGGCACAACCAGAGCCCAGGGCTGAGAGGGAAACCTCATCCAGCAGACGGAGGAGA
CACCCCGATGAGCCTGGACAGTTTGCTGTTGCCCTGGTTGTCATTTACCGGACCCTGGG
TCAGCTGCTGCCTGAACACTATGACCCCGACCATCGCAGCCTCCGACTGCCTAACCGGC
CTGTCATCAACACCCCCGTGGTGAGTGCTATGGTGTACAGTGAGGGAACCCCACTCCCC
AGCTCTCTGCAGAGGCCTATCCTGGTGGAGTTCTCCCTGTTGGAGACGGAGGAACGAAG
CAAACCTGTCTGTGTATTCTGGAACCACTCCCTCGACACTGGTGGGACTGGAGGGTGGT
CAGCCAAGGGCTGTGAACTTCTGTCGAGGAACCGGACCCACGTCACTTGCCAGTGCAGC
CATTCGGCCAGCTGCGCGGTGCTCATGGACATTTCCAGACGTGAGCACGGGGAGGTTCT

FIGURE 9 CONTINUED

GCCCCTGAAGATCATCACCTATGCCGCCCTGTCCTTGTCTTTGGTGGCCCTCCTGGTGG
CCTTTGTCCTTCTCTCGCTCGTTCGGACACTGCGCTCCAACCTGCACAGCATCCCACAA
GAACCTATCCACGCTCTGTTCTTCTCCCAGCTCATCTTCATGGTCGGCATCAACCAGAC
TGAGAACCCGTTTCTCTGCACAGTGGTCGCCATCCTCCTGCACTACGTCTCCATGGGCA
CCTTCGCCTGGACCCTTGTGGAGAACTTGCATGTCTACCGCATGCTGACAGAAGTGCGC
AACATCGACACTGGGCCCATGGCGTTCTACCACGTGGTGGGCTGGGGCATCCCTGCCAT
TGTCACAGGACTGGCTGTTGGCTTGGACCCTCAGGGCTATGGAAACCCTGACTTCTGCT
GGCTGTCCCTTCAGGATACCCTGATTTGGAGCTTTGCTGGGCCTGTCGGAACGGTTATA
ATCATCAACACAGTCATCTTTGTCCTGTCTGCAAAGGTTTCCTGCCAAAGAAAGCACCA
TTATTATGAAAGAAAGGGGGTTGTCTCCATGCTGAGGACGGCCTTCCTCCTGCTGCTGC
TCGTCACTGCCACCTGGCTGCTGGGACTGCTGGCGGTCAACAGTGACACTCTTAGCTTT
CACTACCTCTTTGCTGCCTTCAGCTGCTTGCAGGGCATCTTTGTCCTCCTGTTCTACTG
CGTGGCCAACAGGGAGGTGCGGAAGCACCTGAGGGCGGTGCTGGCAGGGAAGAAGCTGC
AGCTGGATGACTCGGCCACCACTCGGGCCACTCTGCTAACGCGCTCCCTCAACTGCAAC
AACACCTACAGCGAAGGGTCCAGACATGCTCCGCACCGCCCTGGGCAGTCCACAGCCTC
TCTGGACAGTACCACCAGGGATGAAGGGGTCCAGAAACTCAGTGTGTCCTCTGGCCCAG
CCCGTGGTAACCATGGAGAACCAGATGCATCCTTCATCCCTAGGAACTCCAAAAAAGCT
CACGGCCCTGACTCTGACTCTGACAGTGAGCTGTCCCTGGACGAGCACAGTAGTTCCTA
CGCCTCTTCACACACATCGGACAGCGAGGATGATGGCGGAGAGGCTGAAGACAAATGGA
ATCCGGCTGGGGGCCCCGCCCATAGCACCCCAAAAGCAGATGCTCTGGCCAACCACGTC
CCAGCTGGCTGGCCCGACGAGAGCCTGGCTGGGAGTGACAGTGAGGAGTTGGACACTGA
GCCCCACCTGAAGGTGAGACCAAGGTCAGCGTGGAGTTACACCGGCAGGCGCAGGGCAA
TCACTGTGGCGACCGGCCCTCTGACCCGGAAAGTGGGGTCCTGGCCAAGCCAGTGGCCG
TGCTTAGCAGCCAGCCCCAGGAGCAGCGGAAAGGCATCCTGAAAAACAAAGTCACCTAC
CCGCCGGCCATTGCCAGAGCAGCCACTGAAGTCCCGGCTGCGAGAGAAGCTGGCTGATT
GTGAGCAGAGCCCCACATCCTCCCGCACATCCTCCCTTGGCTCTGGCGATGGTGTCCAT
GCCACTGACTGTGTCATTACCATCAAGACTCCGAGGAGGGAGCCAGGCCGTGAGCATCT
CAATGGGGTGGCCATGAATGTACGCACAGGGAGTGCCAGGCCAACGGTTCTGACTCAGA
GAAACCATGAGGCACAGTCAACCCCACAGACTGCCGGTCAAGCCCTCAGACCTTGAAGC
CTGCCTGGGACTGCTGCCTATGGGACAAGCAGGCACCTTGTGTGAGGGTCCCTGCCATA
GCAGCTGGCTCTACGCAGACCGTCCAGACGGGAAGCCCTTGACCCTCATAGGAGCTCAG
GGCCCAGACTCTGACAAAGTGCCAAAGCCACAGATGTCTCCAGAGGGAGACGTGGACTT
CATTAAGGCTGGCATAGCTCCGTCCTTAGACTGAAGACAGAATCCAAACCATGCTGTAC
AAGAGGCCATTGAGCCAGAGCTGGACTTGGTGAATCATTGTACCGGGCCCTTCAACTGT
CCCGCAGGCCTCTCCTTGTGGTACAAGCCCATCACCACCAGCCTAGCGGTGCCTCTGCA
ACGGCAACCCTGGGTTTTAAATTGCTGTCTAAAAATGTAAATAGATATAAATCTCTCCC
TGGACTTGGGAGAAGATGGGAGCTGTGTATGCTTTACACTGCTTTGACTCTGCAGCCAC
TGGAGAGCCATGAAATGGCATCTCACTCTATTGCCCAAGGAAGCTTGCACAGTTGTACT
TGAATCTGGAATGAGTCAACTCAGCTGGTCCCAGTGCCCAGGTAGGGAGCATATGGGCT
GTGAAGTTGACACAGCTTCTGCAGCTGCTTCCGGCCAGGGAGTGTGGGTGTTCCGCTGG
CCTTTGGAGCACACGGCACATGTGCTGCGTATGTGTGCCTCAGCTCCATGGATCCAGGT
CAGGGCTCACCTGTGAGGAGTGGGCGCCATAGCTATGATATGAAACTCTGACCACCCTG
CCACCCCCACAGCCCCCCCCCCCCCCCCGCTCCTGATGCAGTGAGGCGACTCTGGAGCCT
TTCCCAGTCAGCCCAGTGCGCAGCGGGAGGACTTCTGCTCCTGCTCCCATAAGCTCTCA
GAGTGCTCATGTTCCTGATTCTGAGGGAGCCCAGAGGTCTTCCAACGCTGTCACGGTCC
TTCCTGCCAATGTTGTACTTTTTCAGACTCCAGGGCTCTCTAGGCAGGCAGGTGCTCCA
CCCCTGAAATCTGACCAATGACATCACTTTGCTTCAAATGACCAATTGTGCAAAGAAAC
AAAGCCAGAGTGCCCGTTTTCAATGGTTACCACATTCTTTTTGGAATCTCACACCAAGG
ACCTGTGACCAGCCACTGAGAGCCACGGTGCAGCCAAGGCCAAGGGATGGGAGCCTGGA
GTTACTCAGACACGTTTACTTCAGCATGGACTGTTGTCTGAATCAGGTCCCCAAAGTAC
ATGGGGTGCACAGTCGCTCGGAATGGAGAACCTCAGGCGAGGCGGTCAAAGGCCAGGAC
TTCTCACCAAGCTGTGCTCTCAGAAGTGCACAGGACTGCTAGTGACTGACTGGTGAGAT
GAAAGCCTACAAGACATGGCCTGGGGTCACGCACTGCCAAGAGGCTGACGGGAGGCCCA
GGTAGCCCAAGGATGGCAAGGATACAGAGTGACCTAGCACAGGGGAGCTTCAGTCCCAG
GTGGTACAGCACCGTGACAACTCCGCAACCCACGCCACCTCAGAAGGTGAAGTTTTTGA
TTCGATCACAACTATTAGCAAAACAAACCCTGTCAGTTTTAAACTGTTTTTCTGACCTA
AGACTTTCCTTGACCGTAAGACATGGAGATTTTAACAGGTGTATTTATTACTGTTCAGC

FIGURE 9 CONTINUED

ACTGGATGGCAACACAGGTGAGATGATGCATCTATAATAAATTAAGATTTTTGGATTTG
TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 9 CONTINUED

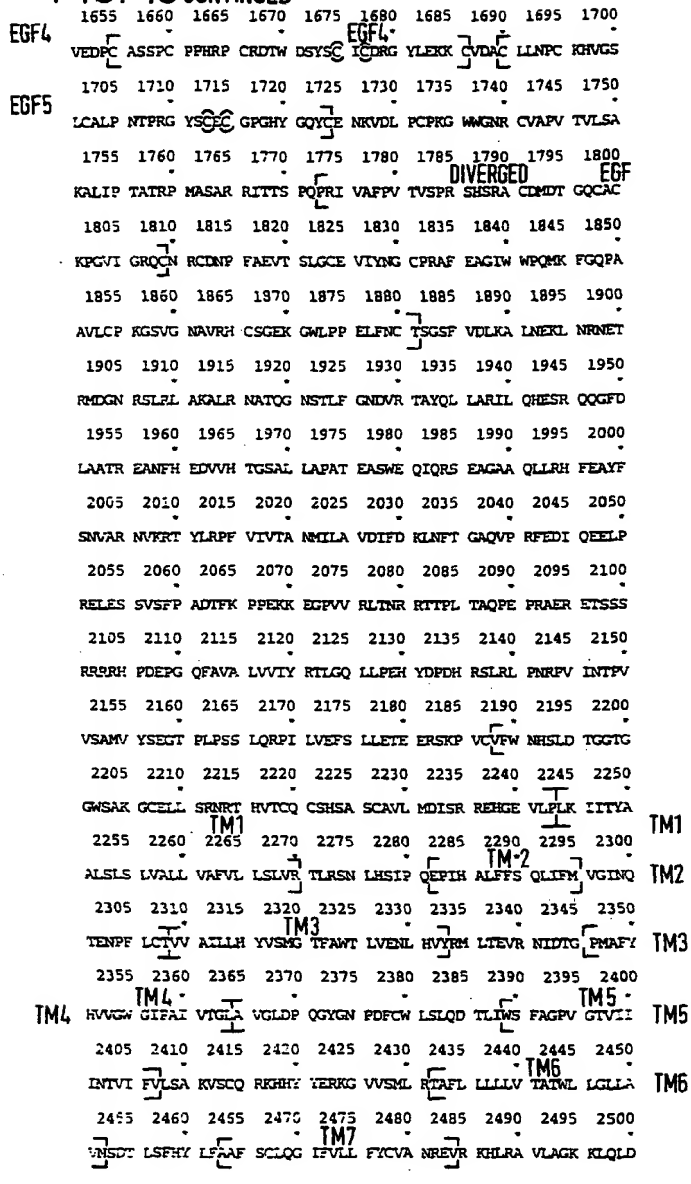
FIG. 10

```
          5    10   15   20   25   30   35   40   45   50
          *    *    *    *    *    *    *    *    *    *
          NSGLM EKLKQ IEEQT KKAQQ ELEEQ TRRAL ELEQE RKRAG TAVIE LRAHD

          55   60   65   70   75   80   85   90   95   100
          *    *    *    *    *    *    *    *    *    *
CD1       PDEGD AGRLS YQMEA LFDER SNGYF LIDAA TGAVT TARSL DRETK DTHVL

          105  110  115  120  125  130  135  140  145  150
          *    *    *    *    *    *    *    *    *    *
          KVSAV DHGSP RRSAA TYLTV TVSDT NDHSP VFEQS EYRER IRENL EVGYE

          155  160  165  170  175  180  185  190  195  200
          *    *    *    *    *    *    *    *    *    *
CD2       VLTIR ATDGD APSNA NMRYR LLEGA GGVFE IDARS GVVRT RAVVD REEAA

          205  210  215  220  225  230  235  240  245  250
          *    *    *    *    *    *    *    *    *    *
          EYQLL VEAND QGRNP GPLSA SATVH IVVED ENDNY PQFSE KRYVV QVPED

          255  260  265  270  275  280  285  290  295  300
          *    *    *    *    *    *    *    *    *    *
CD3       VAVNT AVLRV QATDR DQGQN AAIHY SIVSG NLKGQ FYLHS LSGSL DVINP

          305  310  315  320  325  330  335  340  345  350
          *    *    *    *    *    *    *    *    *    *
          LDFEA IREYT LRIKA QDGGR PPLIN SSGLV SVQVL DVNDN APIFV SSPFQ

          355  360  365  370  375  380  385  390  395  400
          *    *    *    *    *    *    *    *    *    *
          AAVLE NVPLG HSVLH IQAVD ADAGE NARLQ YRLVD TASTI VGGSS VDSEN

CD4       405  410  415  420  425  430  435  440  445  450
          *    *    *    *    *    *    *    *    *    *
          PASAP DFPFQ IHNSS GWITV CAELD REEVE HYSFG VEAVD HGSPA MSSSA

          455  460  465  470  475  480  485  490  495  500
          *    *    *    *    *    *    *    *    *    *
          SVSIT VLDVN DNDPM FTQPV YELRL NEDAA VGSSV LTLRA RDRDA NSVIT

          505  510  515  520  525  530  535  540  545  550
          *    *    *    *    *    *    *    *    *    *
CD5       YQLTG GNTRN RFALS SQSGG GLITL ALPLD YKQER QYVLA VTASD GTRSH

          555  560  565  570  575  580  585  590  595  600
          *    *    *    *    *    *    *    *    *    *
          TAQVF INVTD ANTHR PVFQS SHYTV SVSED RPVGT SIATI SATDE DTGEN

          605  610  615  620  625  630  635  640  645  650
          *    *    *    *    *    *    *    *    *    *
CD6       ARITY VLEDP VPQFR IDPDT GTIYT MTELD YEDQA AYTLA ITAQD NGIPQ

          655  660  665  670  675  680  685  690  695  700
          *    *    *    *    *    *    *    *    *    *
          KSDTT SLEIL ILDAN DNAPR FLRDF YQGSV FEDAP PSTSV LQVSA TDRDS

          705  710  715  720  725  730  735  740  745  750
          *    *    *    *    *    *    *    *    *    *
CD7       GPNGR LLYTF QGGDD GDGDF YIEPT SGVIR TQRRL DRENV AVYNL WALAV

          755  760  765  770  775  780  785  790  795  800
          *    *    *    *    *    *    *    *    *    *
          DRGSP NPLSA SVGIQ VSVLD INDNP PVFEK DELEL FVEEN SPVGS VVARI
```

FIG. 10 CONTINUED

```
                805   810   815   820   825   830   835   840   845   850
CD8  RANDP DEGPN AQIIY QIVEG NVPEV FQLDL LSGDL RALVE LDFEV RRDYM
                855   860   865   870   875   880   885   890   895   900
     LVVQA TSAPL VSRAT VHIRL LDQND NPPEL PDFQI LFNNY VTNKS NSFPS
                905   910   915   920   925   930   935   940   945   950
CD9  GVIGR IPAAD PDLSD SLNYT FLQGN ELSLL LLDPA TGELQ LSRDL DNNRP
                955   960   965   970   975   980   985   990   995  1000
     LEALM EVSVS DGIHS VTALC TLRVT IITDD MLTNS ITVRL ENMSQ EKFLS
               1005  1010  1015  1020  1025  1030  1035  1040  1045  1050
     PLLSL FVEGV ATVLS TTKDD IFVFN IQNDT DVSSN ILNVT FSALL PGGTR
               1055  1060  1065  1070  1075  1080  1085  1090  1095  1100
     GRFFP SEDLQ EQIYL NRTLL TTISA QRVLP FDDNI CLREP CENYM KCVSV
               1105  1110  1115  1120  1125  1130  1135  1140  1145  1150
     LRFDS SAPFI SSTIV LFRPI HPITG LRCRC PPGFT GDYCE TEIDL CYSNP   EGF1
               1155  1160  1165  1170  1175  1180  1185  1190  1195  1200
     CGANG GCRSR EGGYT CECFE DFIGE HCQVN VRSGR CASGV CKNGG TCVNL   EGF2
               1205  1210  1215  1220  1225  1230  1235  1240  1245  1250
     LIGGF HCVCP PGEYE HPYCE VSTRS FPPQS FVTFR GLRQR FHFTV SLAFA
               1255  1260  1265  1270  1275  1280  1285  1290  1295  1300
     TQDRN ALLLY NGRFN EKHDF IALEI VEEQL QLTFS AGETT TIVTP QVPGG
               1305  1310  1315  1320  1325  1330  1335  1340  1345  1350
     VSDGR WHSVL VQYYN KPNIG HLGLP HGPSG EKVAV VTVDD CDAAV AVHFG
               1355  1360  1365  1370  1375  1380  1385  1390  1395  1400
     SYVGN YSCAA QGTQS GSKKS LDLTG PLLLG GVPNL PEDFP VHSRQ FVGCM
               1405  1410  1415  1420  1425  1430  1435  1440  1445  1450
     RNLSI DGRIV DMAAF IANNG TRAGC ASQRN FCDGT SCQNG GTCVN RWNTY
               1455  1460  1465  1470  1475  1480  1485  1490  1495  1500   EGF3
     LCECP LRFGG KNCEQ AMPHP QRFTG ESVVL WSDLD ITISV PWYLG LMFRT
               1505  1510  1515  1520  1525  1530  1535  1540  1545  1550
     RKEDG VLMEA TAGTS SRLHL QILNS YIRFE VSYGP SDVAS MQLSK SRITD
               1555  1560  1565  1570  1575  1580  1585  1590  1595  1600
     GGWHH LLIEL RSAKE GKDIK YLAVM TLDYG MDQST VQIGN QLPGL KMRTI
               1605  1610  1615  1620  1625  1630  1635  1640  1645  1650
     VIGGV TEDKV SVRHG FRGCM QGVRM GESST NIATL NMNDA LKVRV KDGCD
```

FIG. 10 CONTINUED



```
2505  2510  2515  2520  2525  2530  2535  2540  2545  2550
DSATT RATLL TRSLN CNNTY SEGSR HAPHR PGQST ASLDS TTRDE GVQKL

2555  2560  2565  2570  2575  2580  2585  2590  2595  2600
SVSSG PAPGN HGEPD ASFIP RNSKK AHGPD SDSDS ELSLD EHSSS YASSH

2605  2610  2615  2620  2625  2630  2635  2640  2645  2650
TSDSE DDGGE AEDKW NPAGG PAHST PKADA LANHV PAGWP DESLA GSDSE

2655  2660  2665  2670  2675  2680  2685  2690  2695  2700
ELDTE PHLKV RPRSA WSYTG RRRAI TVATG PLTRK VGSWP SQWPC LAASP

2705
RSSGK AS
```

FIG. 10 CONTINUED